



Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Guided color consistency optimization for image mosaicking

Renping Xie, Menghan Xia, Jian Yao*, Li Li

Computer Vision and Remote Sensing (CVRS) Lab, School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei, PR China



ARTICLE INFO

Article history:

Received 30 November 2016
 Received in revised form 15 November 2017
 Accepted 15 November 2017
 Available online 23 November 2017

Keywords:

Image mosaicking
 Color consistency
 Color correction
 Histogram matching

ABSTRACT

This paper studies the problem of color consistency correction for sequential images with diverse color characteristics. Existing algorithms try to adjust all images to minimize color differences among images under a unified energy framework, however, the results are prone to presenting a consistent but unnatural appearance when the color difference between images is large and diverse. In our approach, this problem is addressed effectively by providing a guided initial solution for the global consistency optimization, which avoids converging to a meaningless integrated solution. First of all, to obtain the reliable intensity correspondences in overlapping regions between image pairs, we creatively propose the histogram extreme point matching algorithm which is robust to image geometrical misalignment to some extents. In the absence of the extra reference information, the guided initial solution is learned from the major tone of the original images by searching some image subset as the reference, whose color characteristics will be transferred to the others via the paths of graph analysis. Thus, the final results via global adjustment will take on a consistent color similar to the appearance of the reference image subset. Several groups of convincing experiments on both the synthetic dataset and the challenging real ones sufficiently demonstrate that the proposed approach can achieve as good or even better results compared with the state-of-the-art approaches.

© 2017 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction

Nowadays, the development in photography technology makes it easy to obtain massive remote sensing images and digital photographs, which presents both an opportunity and a challenge. Image stitching is an important step in the field of remote sensing and panorama stitching, which merges two or multiple images with overlapping areas into a single composite image as seamless as possible in both geometry and color tone. Different illumination and sensor properties may cause the enormous color and brightness differences among images, which can not be effectively concealed by smoothing transition (Levin et al., 2004; Xiong and Pulli, 2009) and image blending (Perez et al., 2003, 2008, 2011, 2014). In order to generate the higher quality and better accuracy result for many applications such as remote sensing image mosaicking (Kerschner, 2001; Li et al., 2015), panorama roaming and virtual tourism (Brown and Lowe, 2007; Xiong and Pulli, 2010; Tian et al., 2002; Snavely et al., 2006), the research about color consistency becomes ever more and more necessary.

In the remote sensing field, most of the works on solving the tonal difference for multi-view mosaicking are radiometric normalization (or gain compensation) (Canty et al., 2004; Canty and Nielsen, 2008; López et al., 2011), and some other methods often take relatively simple treatments (Wang et al., 2005; Li et al., 2015). These models often globally and symmetrically found a gain correction that minimizes color differences over corresponding overlaps, but cannot eliminate tonal difference between two adjacency images to the utmost extent. To overcome this problem, Pan et al. (2010) proposed a global-to-local strategy in which the global processing was based on a linear model and the local optimization was carried out by a nonlinear model which divided each overlap into subareas and performed the linear adjustment in each subarea. Vallet and Lelégard (2013) utilized partial iterates to symmetrize the non-parametric color correction, which simultaneously adjusts two images without preserving one image. Because of only symmetrizing the color correction of an image pair, it is required to iterate the process for mosaic correction. Although these algorithms work for some cases, they may fail to completely compensate for color difference between different views when the lighting conditions vary dramatically. Panoramic stitching can obtain an image with a large field of view and present a broader scene, which is quite popular among landscape, cityscape and

* Corresponding author.

E-mail address: jian.yao@whu.edu.cn (J. Yao).
 URL: <http://cvrs.whu.edu.cn/> (J. Yao).

architectural photographers both at home and abroad. The basic steps of image mosaicking are comprised of geometric alignment, tone correction, seamline searching, and feathering or blending. When there exist less color difference between images, there are many blending algorithms can be directly utilized to effectively eliminate artifacts, such as feathering (Prados et al., 2013), alpha blending (Baudisch and Gutwin, 2004), Poisson blending (Pérez et al., 2003) and Gaussian blending (Popovic et al., 2013). However, enormous color difference among images still generated obvious stitching artifacts in the last stitched or mosaicked panorama if we only utilize image blending to correct color consistence. In image processing and computer graphics communities, lots of color manipulation methods have been developed in recent years. These approaches can be divided into two categories: parametric and non-parametric at a high level (Xu and Mulligan, 2010).

Approaches based on the transformation models are parametric, which assume the color relation between images can be described by a transform: $\hat{\mathbf{I}} = f(\mathbf{I})$, where f stands for any transformation equation for a color vector. Reinhard et al. (2001) were pioneers in establishing the concept of color transfer with an approach to modify the color distribution of the original image based on the global color statistics of an example image in the decorrelated $\alpha\beta$ color space. Their work has been widely used as the baseline by other approaches. To operate the color in the RGB color space directly, Ilie and Welch (2005) utilized a general polynomial transformation to correct the color vector in the RGB color space. Huang and Chen (2009) regarded the RGB values and the 2D coordinates as a feature vector of landmark pixels and built the corresponding relation about landmark pixels between the original image and the target one with the Mahalanobis distance. Naim and Isa (2012) proposed a new method of the 3D distribution rotation which was applied on the 2D two-color channel plane (i.e., the red-green plane, the red-blue one, and the green-blue one) instead of the 3D RGB color model. Exposure compensation and vignetting correction are also the prime technologies that address the color balancing problem when the inputs are partially overlapped images (Goldman, 2010). Through statistical analysis between the mean and the standard deviation of an image, some inherent differences between the low-contrast scene image and the normal one can be found, which provide the accurate contrast restoration of color images (Oakley and Bu, 2007). When there exists a great color difference between the images, the forecast about the level of “airlight” may perform not well. Kim and Pollefeys (2008) utilized corresponding pixels to estimate the exposures and vignetting. This method is robust to noise and outliers. Furthermore, based on the spatial color discrepancy model and the temporal variation model, Shao et al. (2010) introduced a new color correction method with a lower computational complexity and a higher accuracy, which was applied to multi-view images and videos. In order to get a global correct relation, Rizzi et al. (2003) proposed a computational model of the human visual system to adjust color consistency, which is based on the global equalization mechanisms that are “Gray World” and “White Patch”. Tai et al. (2005) proposed a local color transfer scheme based on probabilistic image segmentation and region mapping using the Gaussian mixture models (GMM) and the expectation-maximization (EM) algorithm. Xiang et al. (2009) improved this work in the case that multiple source images are available for selection. The bin-ratio-based histogram distance was used in (Hu et al., 2014), which is more robust to partial matching and histogram normalization. Park et al. (2016) adopted a global color correction model based on a low-rank matrix factorization approach to automatically optimize color consistency. This approach is much more efficient in calculations and robust to outliers for a large number of images. However, there are still obvious color difference in the last stitched image, because the

parametric method always utilizes the stable model to address color consistency. At the same time, the compensation obtained from this stable model cannot entirely eliminate color difference for the whole images.

Non-parametric methods mostly study the overlapped area of two images which consist of the same scene or object and can be utilized to design the color mapping. This mapping can be used to maintain the frame-to-frame consistency. The look-up table method is widely used to record the mapping of the full range of color levels directly. To simplify the look-up table, Yoo et al. (2013) proposed to search the major colors in both the original image and the example one, which cluster and then build the mapping relation between major colors through a defined similarity metric. Actually, the look-up table is often replaced by some kind of curves, like Gamma curve, S-curve and B-splines, etc. Moulon et al. (2013) utilized the intensity values in the quantiles of histograms to depict the mapping relation which were then optimized globally as a convex problem. In order to deal with color histograms as much as possible, Papadakis et al. (2011) proposed a variational formulation to transform two or more color images. Pitié et al. (2005, 2007) presented a transfer method of the N -dimensional probability density function (PDF) to reduce the high-dimensional PDF matching to the one-dimensional PDF matching by Radon transform. Another popular color transfer framework is based on iterative optimization, which generally belongs to one of nonlinear and non-parametric methods (Tehrani et al., 2010; Moulon et al., 2013; Hwang et al., 2014). HaCohen et al. (2011) presented a new framework of image iteration based on fitting a global non-linear parametric color model. To achieve the color consistency, HaCohen et al. (2013) found the minimum of a quadratic cost function by global optimization which includes regularization terms and constraints. Frigo et al. (2011) proposed an example-based Chromatic Adaption Transform (CAT) to obtain the illumination matching and select the dominant colors as optimal mapping between input and example images.

Although the above mentioned approaches have solved some key problems in color transfer effectively and made practical improvements, they only considered the processing for two adjacent images, which is not suitable for the global processing for massive images among which there exist obvious color differences. The global optimization strategy has many advantages of minimizing color differences for dozens or even thousands of images, but the results are prone to presenting an unnatural appearance. In this paper, the guided initial solution for the global optimization is proposed to solve this problem and achieve the color consistency. There are generally three steps of the proposed method, as described in Fig. 1. Firstly, we convert the input images into the $\alpha\beta$ color space before tonal correction. To figure out the corresponding relationship of gray values in overlapping areas between two adjacent images, we creatively introduce the histogram extreme point matching strategy based on the feature vectors of the histogram peaks. Then, based on a color difference optimization framework for massive images, if we do not supplement other constraints, the solution is mostly meaningless though it can result in a consistent appearance. To address this issue, we search the optimal reference images and figure out the mapping order as the guided initial solution, referring to which other images will be transformed. Finally, a new global optimization framework is proposed to eliminate the subtle color difference after applying the initial solution, which regards the color difference of overlaps between two adjacent images and the distance with the original mapping curve as the data term and the regularity one, respectively. Our approach also allows the user to construct the optimal mapping order by selecting several adjacent images as the maximum-consistent subset. Experimental results on both the

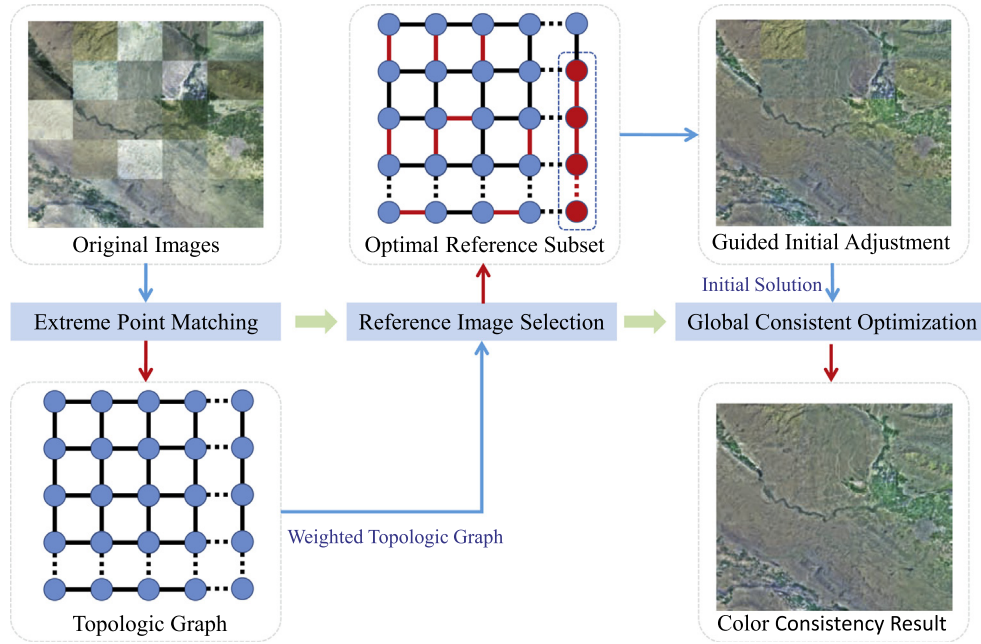


Fig. 1. The flowchart of our proposed color consistency correction framework for image mosaicking.

synthetic dataset and the challenging real ones sufficiently demonstrate that the proposed color consistency strategy for image mosaicking performed better with our method and the main color tone of images to be stitched can be selected to satisfy the user specified constrains when using the user preference editing.

The remainder of this paper is organized as follows. The problem statement about the color mapping model is briefly given in Section 2. Section 3 describes the proposed color correction optimization strategy for image mosaicking in details. The experimental results are reported in Section 4 followed by conclusions and future works drawn in Section 5.

2. Problem statement

Considering the existing correlations between channels in the mostly-used RGB color space, all the pixel operations in our approach are performed in the orthogonal color space $l\alpha\beta$, which makes no requirement of dealing with different channels in a collaborative way. Given a sequence of images with overlaps, our goal is to generate the consistent color of the composited mosaic image by remapping the intensities of the pixels of each image into new values according to its own mapping function. More exactly, the mapping function is defined as three monotonically increasing mapping curves (one per channel), each of which is formulated as a piecewise-quadratic spline with Q control knots ($Q = 6$ was used in this paper). As illustrated in Fig. 2, the coordinates of the unknown knots $\{(v_q, v'_q)\}_{q=1}^Q$ are partly determined by the intensity range of the original image, where $\{(v_q)\}_{q=1}^Q$ are fixed evenly on the horizontal axis to control the mapping curve effectively, while $\{(v'_q)\}_{q=1}^Q$ are free to determine the shape of the mapping curve as the actually unknown parameters. Therefore, the mapping function for an image I can be parameterized as:

$$f = \arg \{(v'_1, \dots, v'_Q)\}_{c=1}^3 \quad (1)$$

where c is the label number for each channel (corresponding to l, α , and β , respectively). That's to say, the degree of freedom of this mapping function is $3Q$.

With the defined mapping model, our goal is to find such a group of mapping curves that can transform the sequential images into a consistent and natural appearance. To solve this problem, we propose to select a set of images with most suitable color characteristics from the original images as the reference subset and find the optimal transferring path by the shortest path algorithm, thought which the corresponding mapping curves can be figured out as the initial solution. Then we perform a global optimization to refine the consistency, instead of taking the risk of creating a consistent but unnatural tone by performing the global optimization directly, as demonstrated in Fig. 3.

3. Our approach

As a problem of non-linear optimization adjusting all the mapping functions' parameters jointly to obtain the ideal result, there are three essential components needed to be considered: the constraint information or measurements, the initial solution, and the energy function. Here, the intensity correspondences in the overlapping regions between images are used as the constraint during energy optimization. The initial solution is determined by transferring the color characteristics of the selected image subset onto others via cascading intensity mapping, and the energy function is comprised of two terms: the differences of the corresponding intensities between images as the data term and the deviation from the initial solution as the penalty one.

3.1. Histogram extreme point matching

Due to some factors such as the illumination variations and different exposure settings, the same objects in the overlaps between adjacent images often present different colors. Even under the same illumination and exposure setting, the same positions possibly present different colors because of the geometric misalignment. To build the color corresponding relationship of the adjacently overlapped images, we select the peaks of a histogram as the feature vectors which combine the information of peaks in Probability Density Function (PDF) and Cumulative Distribution Function (CDF) which make the algorithm more robust. Specifi-

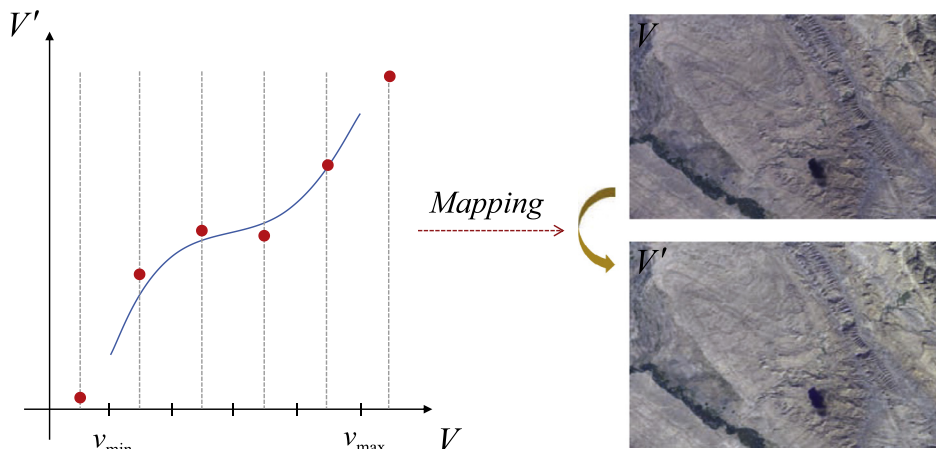


Fig. 2. An illustration of the mapping model used in our approach: (Left) the quadratic spline curve parameterized by 6 knots with their positions evenly fixed on the horizontal (v) axis and free on the vertical (v') axis where v_{\min} and v_{\max} denote the minimal and maximal intensities of the overlapped image region, respectively; (Right) the appearances of an example image before and after applying the intensity mapping.

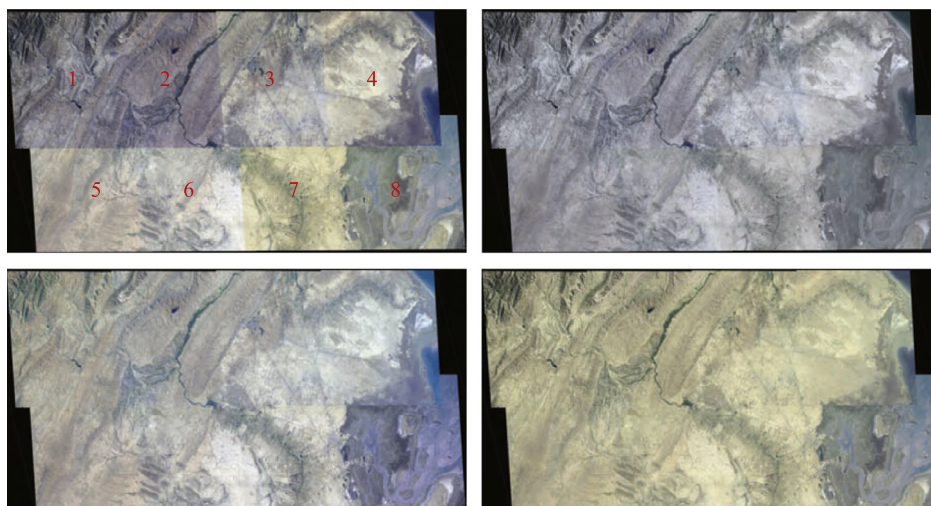


Fig. 3. An illustration of the advantage of the proposed algorithm based on the guided color consistency optimization: (Top-Left) original images (labeled by numbers) captured at different times under different lighting conditions; (Top-Right) the corrected result of our approach without applying the guided initial solution, which results in an unnaturally consistent gray tone; (Bottom-Left) the corrected result of our approach with the user-defined reference subset {5, 6}; (Bottom-Right) the corrected result of our approach with the automatically selected reference subset {7, 8}. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

cally, the histogram of each channel ($1, \alpha, \beta$) is built by counting the number of pixels belonging to every interval which are divided evenly from the image intensity range (default: 300 bins), since the pixel intensity is continuous in the $1\alpha\beta$ color space. Let **A** and **B** be the overlapping regions of two adjacent images, and the relevant PDFs and CDFs are denoted as $\mathbf{PDF}_A, \mathbf{CDF}_A$ and $\mathbf{PDF}_B, \mathbf{CDF}_B$, respectively.

3.1.1. Extracting extreme points

To robustly find extreme points in \mathbf{PDF}_A and \mathbf{PDF}_B , we firstly apply the Gaussian filter on them to suppress possible noise. According to the common definition of the histogram peak, the initial local extreme points can be easily obtained from the filtered \mathbf{PDF}_A and \mathbf{PDF}_B . To address the problem that extreme points may be centralized locally, we further check out all initial extreme points by a local window suppression. Let $\{v_A^i\}_{i=1}^m$ be the intensities of m extreme points $\{P_A^i\}_{i=1}^m$ in \mathbf{PDF}_A , which are sorted in the ascending order. Given an extreme point P_A^i , we select the extreme point(s) within the neighborhood $[v_A^i - w, v_A^i + w]$ centered on the

corresponding intensity v_A^i with the size of $(2w + 1)$ ($w = 2$ in default), and only retain the extreme point with the highest frequency in \mathbf{PDF}_A when there exist multiple extreme points in that neighborhood. Similarly, all initial extreme points are checked in this way. Finally, we obtain the final extreme points of \mathbf{PDF}_A and \mathbf{PDF}_B , which are represented as $\{P_A^i\}_{i=1}^{M_A}$ and $\{P_B^j\}_{j=1}^{M_B}$, respectively, where M_A and M_B are the numbers of the retained extreme points in \mathbf{PDF}_A and \mathbf{PDF}_B , respectively. For each extreme point **P**, its feature vector consists of 4 components $\{F, v, \hat{C}, \check{C}\}$ where F denotes the frequency of this point in PDF, v stands for the corresponding intensity, and \hat{C} and \check{C} denote the cumulative values of intensities $(v - \epsilon)$ and $(v + \epsilon)$ in CDF ($\epsilon = 2$ in default).

3.1.2. Matching extreme points

The extreme points can reflect image statistical characteristics in considerable degree. Making the intensities of corresponding extreme points as nearby as possible is an effective way to decrease the color differences between images. To reliably match

these extreme points $\{\mathbf{P}_A^i\}_{i=1}^{M_A}$ and $\{\mathbf{P}_B^j\}_{j=1}^{M_B}$, we define a cost function to measure the matching similarity of two extreme points \mathbf{P}_A^i and \mathbf{P}_B^j as:

$$\text{Cost}(\mathbf{P}_A^i, \mathbf{P}_B^j) = \frac{F_A^i + F_B^j}{2F_{\max}} \times \frac{\min(F_A^i, F_B^j)}{\max(F_A^i, F_B^j)} \times \frac{\max(\hat{C}_A^i - \check{C}_A^i, \hat{C}_B^j - \check{C}_B^j)}{\max(\hat{C}_A^i, \hat{C}_B^j) - \min(\check{C}_A^i, \check{C}_B^j)}, \quad (2)$$

where F_{\max} is the maximal frequency of all the extreme points in PDF_A and PDF_B . From the above equation, we can find that when the frequencies of an extreme point pair are larger and more similar, and their accumulative values are more nearby, their matching cost is bigger. The higher the cost function value is, the more likely the two points are matched. Based on this cost definition, a matching cost matrix $\mathbf{W} = [W_{ij}]_{M_A \times M_B}$ is created. In order to efficiently remove the impossibly matched extreme points, we design three hard conditions from the views of both PDF and CDF to further examine \mathbf{P}_A^i and \mathbf{P}_B^j according to:

$$\begin{cases} \frac{\min(F_A^i, F_B^j)}{\max(F_A^i, F_B^j)} < \theta_f, \\ \hat{C}_A^i > \check{C}_B^j + \theta_c \times C_{\max}, \\ \hat{C}_B^j > \check{C}_A^i + \theta_c \times C_{\max}, \end{cases} \quad (3)$$

where θ_f and θ_c are two given thresholds, and C_{\max} is the maximal value of CDF, which means the valid pixel number of overlapping regions. The first term $\frac{\min(F_A^i, F_B^j)}{\max(F_A^i, F_B^j)}$ indicates that the similar frequencies

for two possibly matched extreme points \mathbf{P}_A^i and \mathbf{P}_B^j in PDF. The lower the value of θ_f we set is, the allowable difference of frequencies between two extreme points is bigger. We set $\theta_f = 0.25$ by default, which is used to measure the distance between two extreme points \mathbf{P}_A^i and \mathbf{P}_B^j in CDF. The higher the value of θ_c is, the allowable distance between two extreme points in CDF is bigger. We set $\theta_c = 0.02$ by default. If at least one of the above three conditions is met, which indicates that \mathbf{P}_A^i and \mathbf{P}_B^j are not possibly matched, their matching cost is set to zero, i.e., $\mathbf{W}_{ij} = \text{Cost}(\mathbf{P}_A^i, \mathbf{P}_B^j) = 0$. From the view of PDF, the first condition indicates that the frequencies of the two possibly matched extreme points \mathbf{P}_A^i and \mathbf{P}_B^j should be a relatively small difference. From the view of CDF, the second and third conditions indicate that \mathbf{P}_A^i and \mathbf{P}_B^j are more likely to be matched when their corresponding CDF values are approximate. According to the above three hard conditions, the matching cost matrix \mathbf{W} will be updated, in which all the zero elements indicate that they are not possibly matched.

Based on the computed matching cost matrix \mathbf{W} , we propose an efficient iterative strategy to find the matched extreme points as:

- **Step 1:** Finding the highest non-zero cost element \mathbf{W}_{ij} from the matrix \mathbf{W} and its corresponding extreme points \mathbf{P}_A^i and \mathbf{P}_B^j are selected out as a reliable extreme point match.
- **Step 2:** Updating the matrix \mathbf{W} by removing the i -th row and the j -th column due to that \mathbf{P}_A^i and \mathbf{P}_B^j have been matched successfully.

By performing the above two steps iteratively until the updated matrix \mathbf{W} is empty or there exists no non-zero element in \mathbf{W} , a set of reliable extreme point matches will be found, as an example result of extreme points matching shown in Fig. 4. The bin of the horizontal axis is the index of intervals which divide the dynamic intensity range into 300 uniformly spaced samples for each color channel in the $1\alpha\beta$ color space.

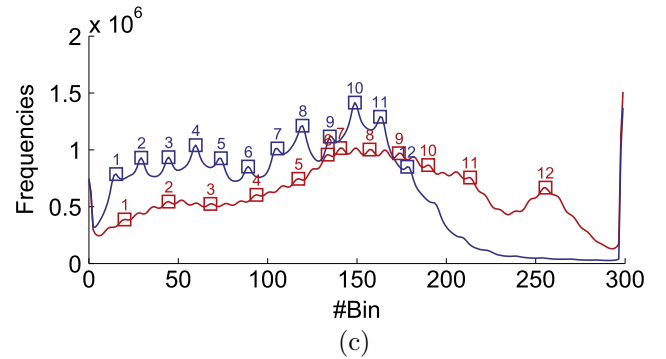
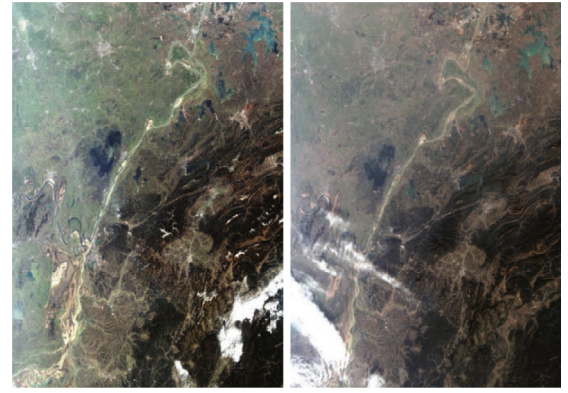


Fig. 4. A visual example of our proposed histogram extreme point matching strategy: (a)-(b) the overlapped image regions of two adjacent images, respectively; (c) the curves of PDF in one channel where the red curve stands for the left image and the blue one stands for the right one. Those points with the same digital labels are regarded to be matched point pairs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Sometimes, no match or too few matches can be reliably found via the above matching strategy in the whole CDF range or some relatively large CDF range. In this case, we will introduce more matches with the help of both CDF_A and CDF_B , which are selected from H uniformly distributed points $\{C_A^k\}_{k=1}^H$ and $\{C_B^k\}_{k=1}^H$ from CDF_A and CDF_B , respectively, but not from the previously found extreme points. The same number of sampling points in CDF_A and CDF_B are uniformly selected in accordance with the cumulative density values. If there exists no extreme point match found in the ranges $[C_A^k - \kappa C_{\max}, C_A^k + \kappa C_{\max}]$ and $[C_B^k - \kappa C_{\max}, C_B^k + \kappa C_{\max}]$, the current sampling points C_A^k and C_B^k will be added into the matching set as a new point match, where κ is a given threshold in advance ($\kappa = 0.1$ in default).

To demonstrate the advantage of our proposed histogram extreme point matching strategy, we conducted the experiment on six image pairs with overlap (the first five pairs from panorama images and the last one from remote sensing ones) via color correction optimization based on our proposed histogram extreme point matching strategy and the widely-used CDF probability correspondence strategy. In general, the CDF probability correspondences between two overlapped images can be built based on the evenly sampled points on their corresponding CDF curves. The histogram distances (i.e., tonal differences) in the l channel of six image pairs of finally corrected results by applying different color correspondence strategies are shown in Fig. 5(e). We observed that our proposed matching strategy can build better correspondences which will result in lower histogram distances. The same conclusions can be observed in the α and β channels. For

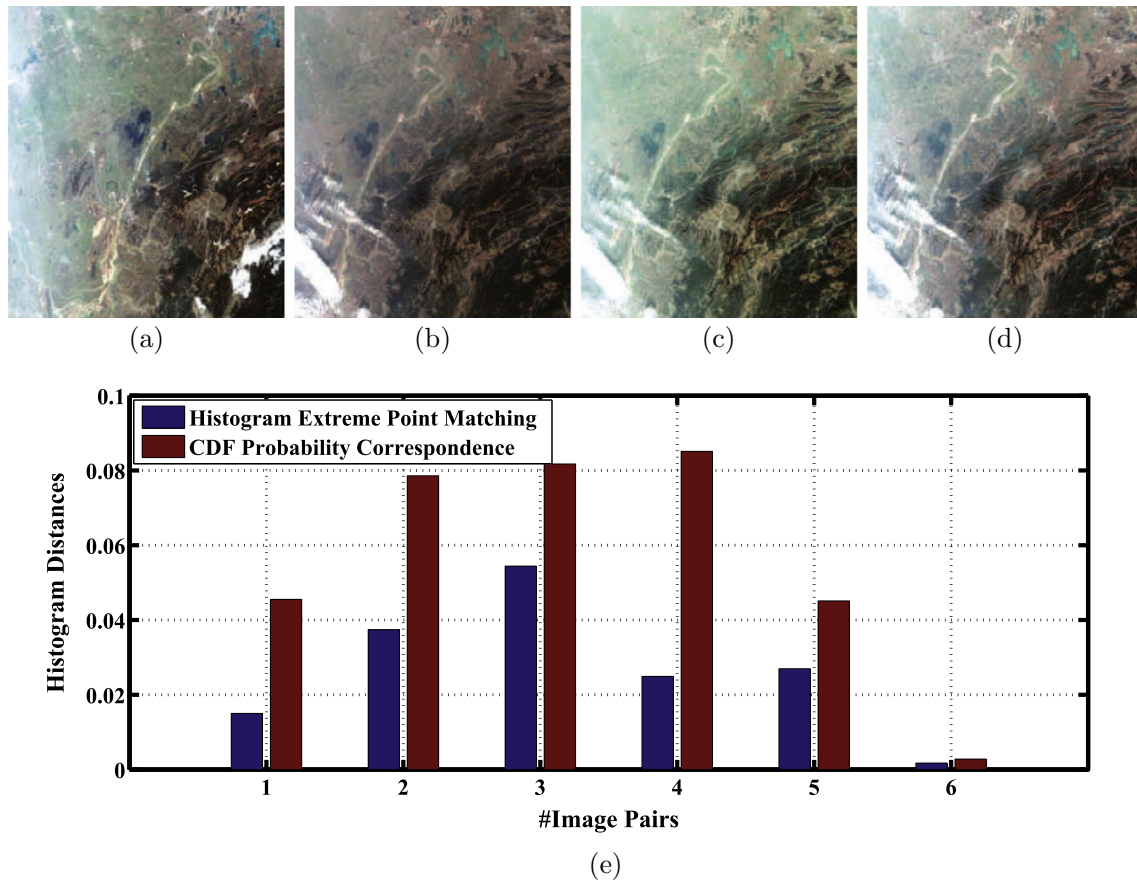


Fig. 5. A comparative illustration of the proposed histogram extreme point matching strategy and the CDF probability correspondence strategy: (a) the reference image; (b) the original image; (c) the result corrected by the CDF probability correspondence strategy; (d) the result corrected by the proposed histogram extreme point matching strategy; (e) the histogram distances (i.e., tonal differences) in the 1 channels by applying two different matching strategies on six image pairs.

visual comparison, we chose the fifth image pair as shown in Fig. 5 (a-b), which are regarded as the original image and the reference one, respectively. The finally corrected results by applying two color correspondence strategies are shown in Fig. 5(c) and (d), respectively. We can find that there exist obvious color deviation caused by incorrect corresponding points in the corrected image, as shown in Fig. 5(c), which was obtained by applying the CDF probability correspondence strategy. While, the image corrected by our approach, as shown in Fig. 5(d), presents a good color consistency with the reference image as shown in Fig. 5(a). In this selected image pair, the nearly fifty percent overlap between two images is covered by green vegetation. In this case, the CDF probability correspondence strategy is prone to building incorrect corresponding point pairs. However, our proposed matching strategy is more robust. The above experiment results sufficiently demonstrate that our proposed matching strategy outperformed the CDF probability correspondence strategy.

3.2. Guided initial solution

As aforementioned, we have to select an image subset as the reference to transfer their color characteristics onto all the other images, whose relevant mapping curves are served as the initial solution for the following global optimization. Without extra reference information available, we adopt a heuristic strategy to take the major tone of the original images as the reference, namely to search the maximum subset of images with consistent color.

3.2.1. Histogram distance

To judge the color difference between two adjacent images, we define the distance between histograms as a quantitative metric. Given an image pair $(\mathbf{I}_i, \mathbf{I}_j)$, we utilize the matched extreme point pair set $\{\mathbf{P}_k^i, \mathbf{P}_k^j\}_{k=1}^{K_{ij}}$ whose corresponding intensity values are $\{v_k^i, v_k^j\}_{k=1}^{K_{ij}}$ to fit a quadratic spline curve in the coordinate system $o v^i v^j$ as an example shown in Fig. 6(a). The distance between two histograms in the overlapping regions between \mathbf{I}_i and \mathbf{I}_j is measured as:

$$d_H(\mathbf{I}_i, \mathbf{I}_j) = A_s / \max(|v_1^i - v_{K_{ij}}^j|, |v_1^j - v_{K_{ij}}^i|), \quad (4)$$

where K_{ij} is the number of matched histogram extreme points, A_s denotes the area of the shadow region formed by the fitted curve and the unit-slope line as shown in Fig. 6(a). It is easy to understand that the fitted curve will coincide with the unit-slope line when there is no difference between the two histograms.

3.2.2. Reference subset

The optimal reference images are selected through two ways: automatic selection and user editing. The detailed descriptions about user preference editing are presented in Section 3.4 and the strategy of automatic reference selection will be discussed below. After histogram distance calculation, we can construct a graph of images in which each image is regarded as a node and the neighboring relationship with overlap is represented as an arc. The cost of an arc linking two images \mathbf{I}_i and \mathbf{I}_j is defined as:

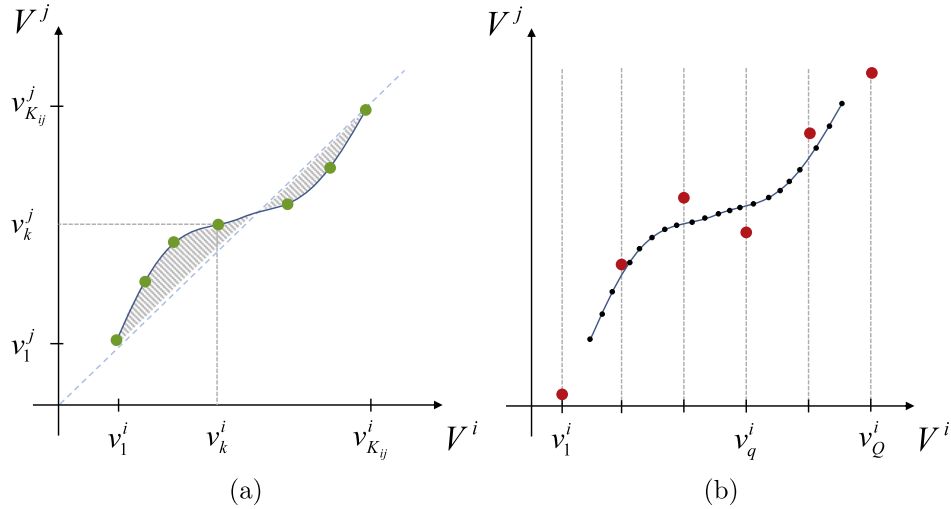


Fig. 6. (a) shadow area formed by the unit-slop line and the quadratic spline curve fitted from the extreme points pairs (labeled as green dots) between \mathbf{I}_i and \mathbf{I}_j ; (b) discrete points (little black dots) picked evenly on the fitted curve, which are used to solve the mapping model parameters of \mathbf{I}_i , namely to determine the positions of the red knots. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$c(\mathbf{I}_i, \mathbf{I}_j) = \frac{1}{3} \left(d_H^\alpha(\mathbf{I}_i, \mathbf{I}_j) + d_H^\beta(\mathbf{I}_i, \mathbf{I}_j) + c_{\text{base}} \right), \quad (5)$$

where $d_H^\alpha(\mathbf{I}_i, \mathbf{I}_j)$ and $d_H^\beta(\mathbf{I}_i, \mathbf{I}_j)$ denote the distances between histograms of the overlapping regions between \mathbf{I}_i and \mathbf{I}_j in the α and β channels, respectively. Note that the cost value won't be affected by the difference in the luminance channel, since the satisfying luminance consistency is easy to be obtained in the global optimization and should not interfere the selection of color reference. Specially, the regularity term c_{base} , set as 1 in our experiments, aims to avoid to find a least-cost path across too many nodes in some extreme cases. According to the cost values that are bigger or smaller than the given thresholds in the α and β channels, the arcs are classified into *consistent arcs* and *inconsistent ones*, as shown in Fig. 7(a). In our experiments, the threshold combination was set as (0.03, 0.005) for the α and β channels, respectively, which can always give the ideal results according to a lots of testing results. As a result, we can get some connected sub graphs with the consistent arcs, for example, the image subsets {08,13,14} and {17,18,19,20} shown in Fig. 7(a), and these image subsets were selected as candidate sets for reference image subset. The nodes of the maximum connected sub graph with the consistent arcs are automatically selected as the reference image subset.

3.2.3. Optimal transferring path

By applying the shortest path algorithm, we can find the optimal transferring path from the root (i.e., the selected reference image subset) to each of other nodes by searching on the spanning tree as illustrated in Fig. 7(b). Thus, the color characteristics of the reference images can be transferred to any other image by cascading the mapping relations along the optimal path, which can be described as a mapping curve fitted by the updated extreme point pairs.

To avoid an extreme case in which there are too few matches, first of all, we fit the mapping curves based on the matched histogram extreme points, and then select some denser samples evenly on these curves. According to the evenly-distributed samples on the mapping curves, we can solve the unknown control points of the mapping model to best fit the samples as shown in Fig. 6(b), which are used as the initial solution $\{\bar{f}_i\}_{i=1}^N$ of the mapping curves $\{f_i\}_{i=1}^N$ for the following global optimization where N is the total number of images.

3.3. Global consistency optimization

Although all the images may present the similar color characteristics to those of the original reference images via applying the initial solution, there probably exist some subtle color differences among images since these deviations have never been processed in a global way. In this section, we propose to adjust all the images jointly in a unified energy framework to increase the global color consistency, and the optimization procedure is conducted independently in each channel with the energy function formulated as:

$$E = \sum_{\mathbf{I}_i \setminus \mathbf{I}_j \neq \emptyset, i, j \in [1, N]} w_{ij} E_{\text{data}}(f_i, f_j) + \lambda \sum_{i=1}^N E_{\text{smooth}}(f_i) \\ \text{subject to: } f'_i > 0, \forall i \in [1, N] \quad (6)$$

where w_{ij} is the weight proportional to the area of the overlapping region between \mathbf{I}_i and \mathbf{I}_j , N is the total number of all the images, and the first-order derivative f'_i of f_i is used to ensure the mapping monotonically increasing. The parameter λ denotes the weight coefficient (we set it as 200 by default) for balancing the data term and the regularity one, which are respectively expressed as:

$$\begin{cases} E_{\text{data}}(f_i, f_j) = \sum_{k=1}^{K_{ij}} (f_i(v_k^i) - f_j(v_k^j))^2, \\ E_{\text{smooth}}(f_i) = \sum_{q=1}^Q (f_i(v_q^i) - \bar{f}_i(v_q^i))^2, \end{cases} \quad (7)$$

where K_{ij} denotes the number of the matched extreme point pairs $\{\mathbf{P}_k^i, \mathbf{P}_k^j\}_{k=1}^{K_{ij}}$ with the intensities $\{v_k^i, v_k^j\}_{k=1}^{K_{ij}}$ between \mathbf{I}_i and \mathbf{I}_j , and Q is the number of evenly-distributed control knots $\{v_q^i\}_{q=1}^Q$, which is the same for all images ($Q = 6$ was used in this paper). Specially, $\{\bar{f}_i\}_{i=1}^N$ stands for the initial solutions of $\{f_i\}_{i=1}^N$, which are computed from the operation detailed in Section 3.2. The above defined data term is used to measure the color difference between adjacent images based on the matched extreme point pairs whose corresponding transformed intensities will be iteratively interpolated from the fitted mapping curves based on the optimized evenly-distributed control knots $\{v_q^i\}_{q=1}^Q$, while the smooth term is used to ensure that the optimized mapping curves $\{f_i\}_{i=1}^N$ do not deviate

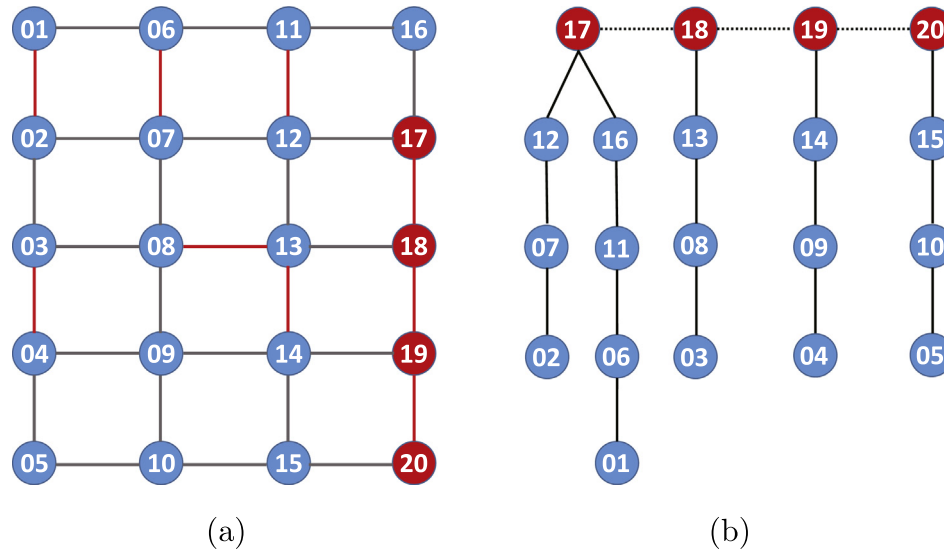


Fig. 7. An illustration of finding the optimal reference subset and the optimal transferring order: (a) the graph of 20 images in 4 strips with 5 images each, whose arcs are labeled in red (consistent arcs) or in blue (inconsistent arcs), where the nodes of the maximum connected sub graph with the red arcs are selected as the reference image subset; (b) the spanning tree generated by the shortest path algorithm, which is applied with the costs between the reference images set as 0. Note: the nodes of the reference subset are labeled in red as the root ones. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

too far from the initial solution $\{\bar{f}_i\}_{i=1}^N$ which represents the tonal characteristic of natural meanings, and the gradient of quadratic spline is increasing monotonicity. The above optimization can be easily solved by the Levenberg–Marquardt (LM) algorithm. Note that the initial solution $\{\bar{f}_i\}_{i=1}^N$ are parameterized with the matched extreme point pairs obtained via the optimal transferring path while the finally optimized solution $\{f_i\}_{i=1}^N$ are parameterized with the evenly-distributed control knots.

It should be noted that the non-uniformly matched extreme point pairs are used as the constrain information or measurements during global optimization and the evenly-distributed control knots are unknown parameters to be optimized in this paper. For two images, we can also regard the matched histogram extreme points as unknown parameters, which can better represent the monolithic color characteristic. Nevertheless, in the multi-view case, each image always has several overlapping neighbors with distinctive tones and we can get multiple sets of matched extreme points for each image, based on which it is hard to adjust the color tones of adjacent images jointly. In order to address this important issue, we select the evenly-distributed control knots on the mapping curves as unknown parameters based on which color tones can be adjusted jointly.

3.4. User preference editing

As described above, our color consistency algorithm is automatic completely and doesn't need the user input or the parameter setting to obtain a consistent color appearance across all the images. To guarantee the result to be as natural as possible, the final color characteristics will be similar to those of the reference images selected during constructing the initial solution for guiding. The reference images are selected based on the principle keeping the major tone of the original images, however, users may prefer some particular color characteristic presented on some original image(s). To provide the function of the user editing, users can select the reference images manually during building the initial solution and the color appearance will propagate across all the images after the process of our algorithm, as an illustration shown in Fig. 3(Bottom-Left).

When there exist obvious color differences among a large set of images covering a wide area, only one local reference image subset cannot make all images presenting similar color characteristics after applying the initial solution and would possibly result in a large propagation cumulative error for images which are far away from the reference image subset. To overcome this problem, our proposed algorithm allows the user to manually select two or multiple separated image subsets as the reference. This can effectively alleviate the propagation cumulative error caused by applying the initial solution which transfers color characteristics of the reference subsets to any other image by cascading the mapping relations along the optimal transferring path. The optimal mapping graph based on two or multiple reference subsets are utilized as the initial solution for global optimization, as illustrated in Fig. 8 from which we observe that a more consistent color can be achieved by manually selecting two image subsets as the reference. Moreover, to make the color appearance more appealing, the user can edit the color of partial images by some softwares like *Photoshop* or techniques in advance and set them as the reference images in our algorithm.

4. Experimental results

In remote sensing community, there are basically two categories methods to adjust the tone of images. The one is radiance calibration, such as the top-of-atmosphere(TOA) correction. Such kind of algorithms, based on some well-defined physically imaging model, are usually used to remove the effects induced by atmosphere condition or sensors' quality for a single image. The other is tonal correction based on physical model, which aims at turning a group of multi-view images into a consistent tone by adjusting their color palette. Thus, they are mainly used in cartographic applications and image mosaicking visualizations. Our proposed method belongs to the latter. In this section, only tonal correction based methods will be compared while radiance calibration based methods will not be discussed. This is because they only works on dataset whose imaging parameters are available, and their performance on multi-view tonal correction can not match those of tonal correction based methods, as Fig. 9 illustrated.

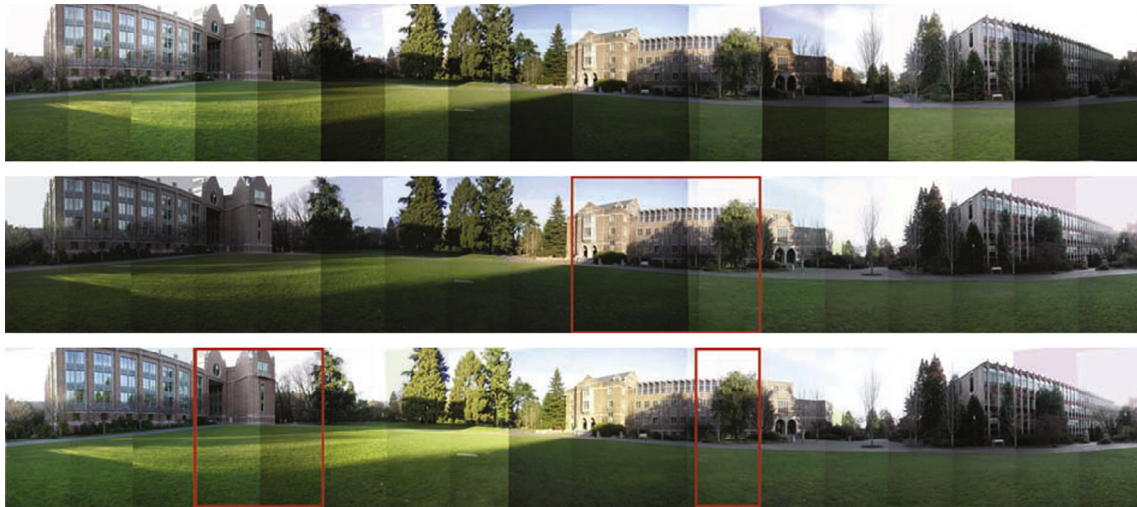


Fig. 8. Comparison after selecting one or two image subsets as reference: (Top) a composited mosaic of original 18 images; (Middle) the mosaicking result after applying the single reference subset marked in the red rectangle; (Bottom) the mosaicking result after applying two reference subsets marked in the red rectangles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

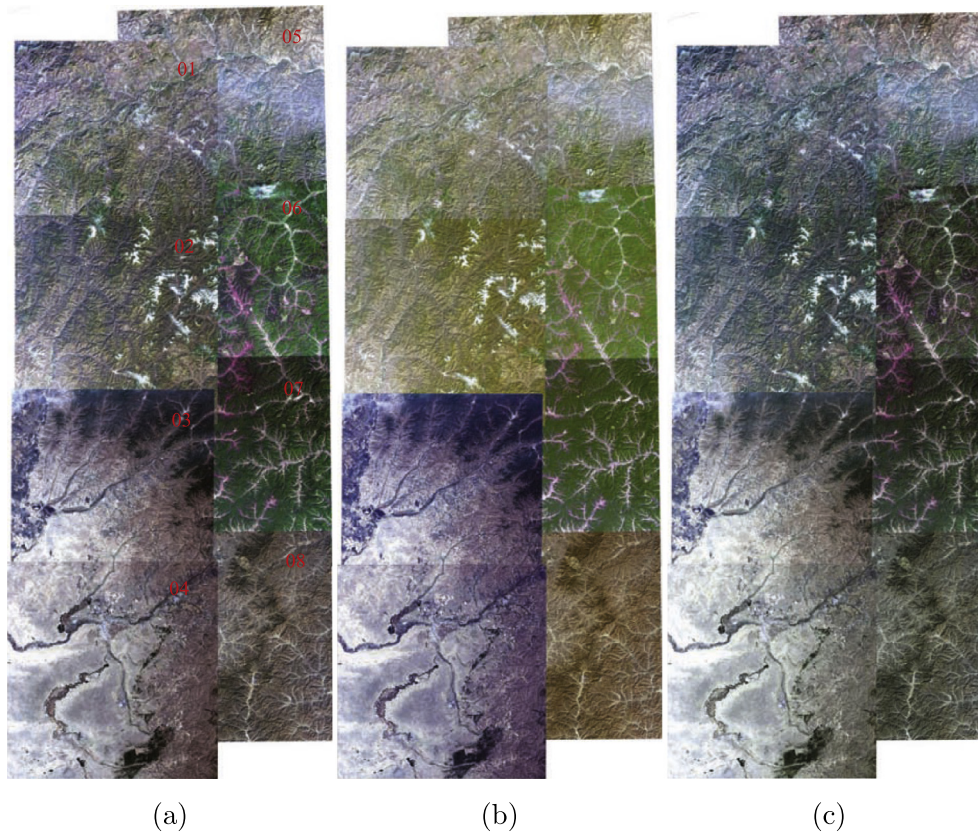


Fig. 9. The experimental results on multispectral images acquired by Landsat8 satellite: (a) original images displayed as a composited mosaic by a simple superposition and labeled by its image index at the upper right corner; (b) the result corrected by TOA correction; (c) the result corrected by our approach with the auto-selected consistent reference image subset {1, 2, 5}. The TOA radiation correction was applied with the FLAASH Atmospheric Correction Module of a commercial software named ENVI on the multispectral images.

To test our method, four typical datasets, including one synthetic dataset and three real ones, were used to evaluate our proposed approach. Along with the experimental analysis, the comparative experiments with the state-of-the-art algorithm proposed by Park et al. (2016) were performed, which were evaluated both qualitatively and quantitatively. The quantitative evaluations

were made by computing the average color difference between all the image pairs in each channel (l, α , and β) according to Eq. (4) and their corresponding consumed time. The comprehensive analysis results demonstrate the superiority of our approach in improving the color consistency of sequential images for mosaicking.

4.1. Synthetic dataset

In this experiment, a synthetic image dataset comprised of 5×5 images was cropped from a wide-view high resolution remote sensing image as shown in Fig. 10(a) with the size of 3676×3127 pixels, in which there exists a 20% overlap between two adjacent images. To test the robustness of our approach, all these cropped images, except for three particularly selected as the reference, were adjusted individually in a commercial software named *Photoshop* to present a diverse appearance, as illustrated in Fig. 10(b). During our experiment, the three unedited images were selected automatically as the optimal reference images in our approach, and they were also kept fixed in the optimization of Park et al.'s approach (Park et al., 2016) to make the comparison meaningful. For the convenience of more detailed comparisons, all the images via different operations are all displayed in the form of the mosaicked image through a simple superposition as shown in Fig. 10(b). From the results shown in Fig. 10(c)–(d), it's easy to find that the color consistency among images via both approaches have been increased so much that there is nearly no difference between the original image and the mosaicked images composited from them. However, compared to Park et al.'s approach, our approach has obtained a more consistent result with the original images, for example, the top right part of the image shown in Fig. 10(c) has not been corrected completely from the color cast using Park et al.'s approach. As a quantitative evaluation, the similarity between each transferred image and its original image was measured by computing the Peak Signal to Noise Ratio (PSNR), and the analysis results illustrated in Fig. 11 show the superiority of

our approach over Park et al.'s approach in keeping the PSNR of the transferred results.

The visual variations on the images shown in Fig. 10(c)–(d) are not obvious enough but the related numerical comparative analysis in each channel shown in Table 1 demonstrates that the color deviation between image pairs greatly decreased after the application of both Park et al.'s approach and ours. Further more, we observe that our approach achieved a more consistent result than Park et al.'s approach in all channels. This benefits from our used mapping model that has more degrees of freedom (DoFs) to adjust the appearance of each image, while Park et al.'s approach only employed a 2-DoF model based on the gamma correction.

4.2. Real multi-temporal datasets

4.2.1. Comparative analysis

Firstly we tested our approach on a set of multi-temporal images with three strips consisting of 21 images with the down-sampling size of 2294×1904 captured by the Chinese satellite ZY-3. These three strips, comprised of 01 ~ 07, 08 ~ 14, and 15 ~ 21 images, respectively, as shown in Fig. 12(a), were captured at different times, 2012/06/26-15:34, 2013/06/15-15:42, and 2012/06/16-15:37, respectively. Because of the different atmosphere conditions or various angles of lighting, there exists the color inconsistency among both intra-strip and inter-strip images. After applying the color consistency optimization, all the images were turned into a similar color characteristic, as shown in Fig. 12(b) and (c) corresponding to Park et al.'s approach and ours, respectively. The final color appearance of the result via our

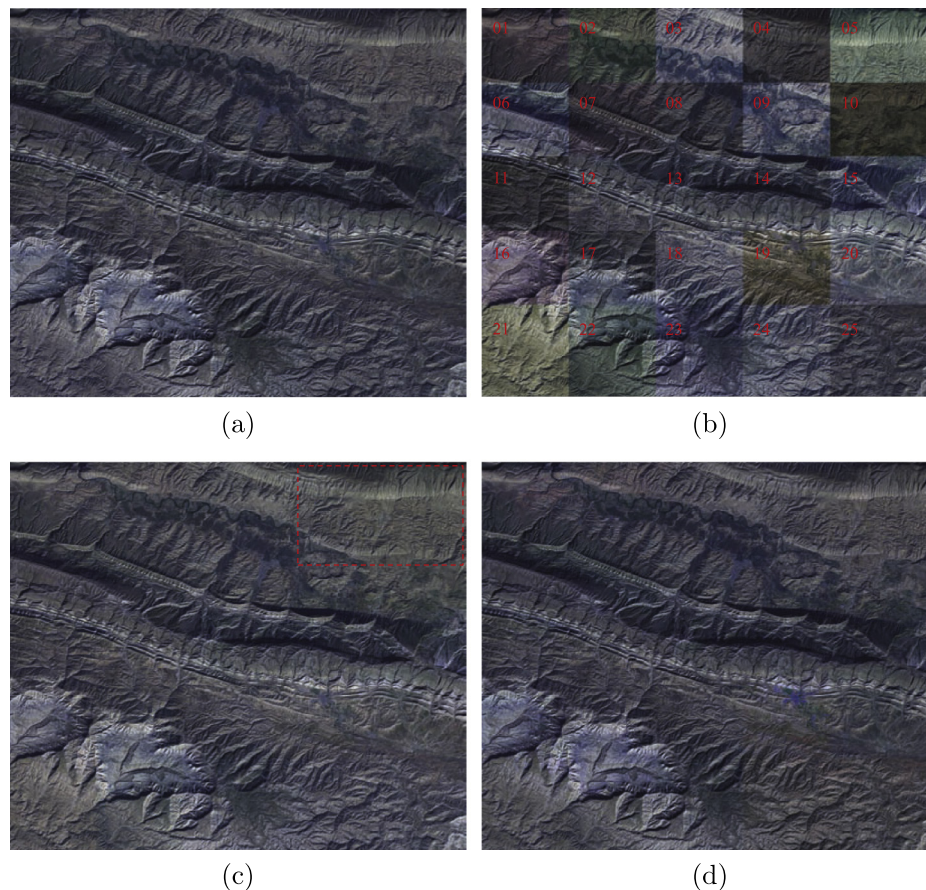


Fig. 10. The experimental results on the synthetic dataset (SYNTHETICS): (a) the original high-resolution image; (b) the cropped images with a 20% overlap adjusted individually via *Photoshop* displayed as a composited mosaicked image; (c) the result corrected by Park et al.'s approach (Park et al., 2016) with image subset {12,13,14} fixed in optimization, and the red box marks an example region inconsistent with the original image in color style; (d) the result corrected by our approach with the auto-selected reference image subset {12,13,14}. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

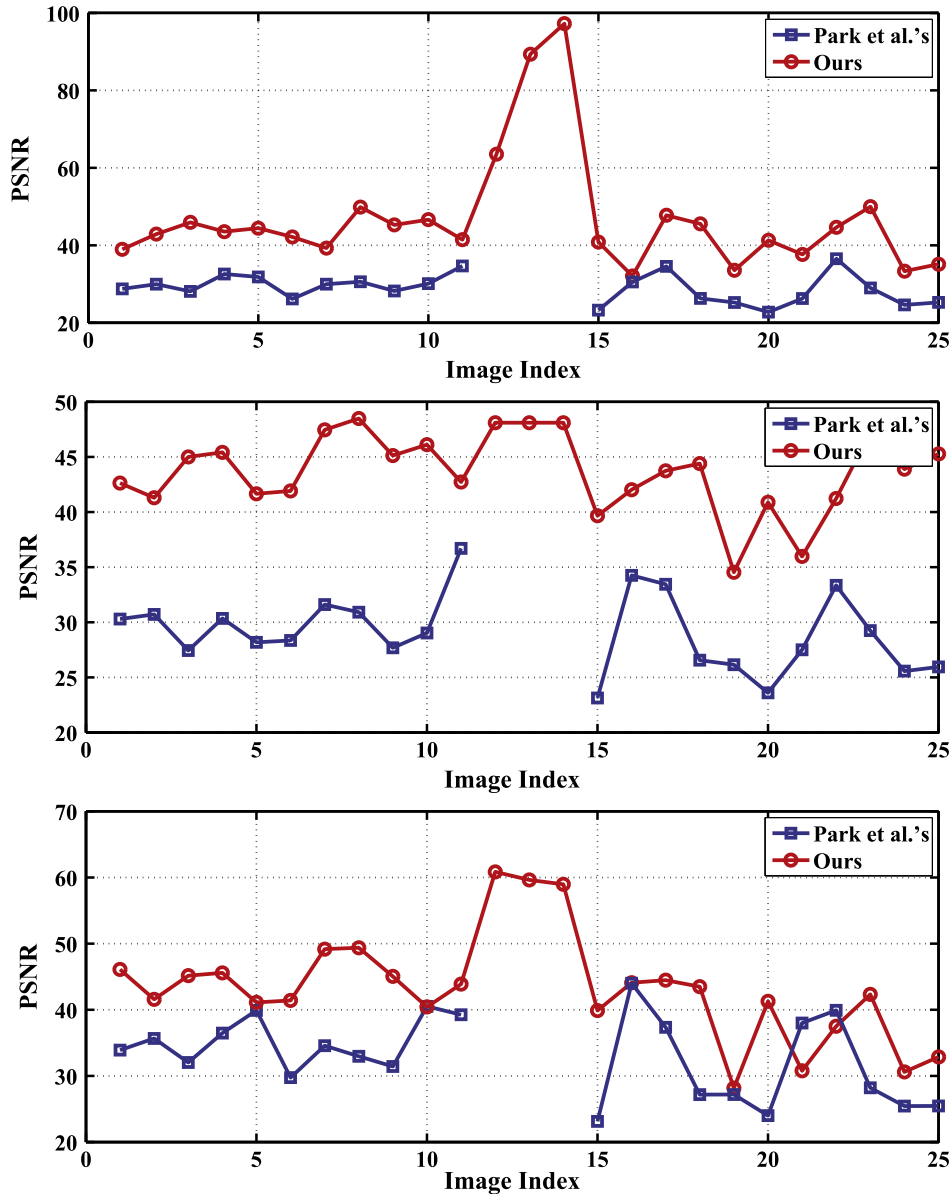


Fig. 11. The PSNR comparative results of the transferred images with respect to their corresponding original ones by Park et al.'s approach and ours, respectively. From top to bottom, the comparative results correspond to the l , α , and β channels, respectively. Note that the values of PSNR for the images {12,13,14} in Park et al.'s approach are omitted since they are totally the same with their original ones.

approach is close to the color characteristics of the auto-selected reference image subset {13,14,19,20,21}. There is a noticeable tonal difference between image strips and residual white tone uncorrected (marked in red boxes) in Fig. 12(c). From the visual comparison, it is hard to say whose result is much better. The numerical comparative analysis for this dataset is reported in Table 1, which demonstrates the superiority of our approach over Park et al.'s in decreasing the color difference among images.

Besides, we should notice that the appearance of the image 07 has not been changed under the operation of Park et al.'s approach. Because of lacking the texture and the low contrast, the feature matching between the image 07 and other images failed, so it has no access to join into the optimization. Therefore, as far as robustness is concerned, our approach relying on the histogram extreme points matching of overlapping regions is superior to Park et al.'s approach relying on feature matching, since the overlapping regions can be determined easily and robustly but sometimes feature matching doesn't work. Moreover, it should be noted that the

quality of the initial solution makes direct effects to the final results, since the finally optimized result is actually a local optimal around the initial solution. For example, a better performance is obtained on the synthetic datasets than dataset-1, and it is because the squared layout of the synthetic data benefits from the central reference image making relatively shorter average transferring path, which tends to provide an initial solution of higher quality for the final optimization.

For a more noticeable comparison on vision, the second multi-temporal dataset consisting of 16 images with a down-sampling size of 2795×2295 , which presents even more severe color difference, was used to test the approaches further. The original images are shown in Fig. 13(a) as a mosaicked image, and the corresponding results of Park et al.'s approach and ours are illustrated in Fig. 13(b) and (c), respectively. The result of Park et al.'s approach has noticeable tonal difference between image strips and residual yellow tone uncorrected (marked in red boxes). Besides, in Fig. 13(b), the mosaicking contour lines between inter-strip images

Table 1
Quantitative evaluations on the experimental results of the four datasets. Unit of time: second (s).

Methods	Item	l	α	β	Time
Synthetics	Input	0.19504	0.02053	0.00370	0.00
	Park et al.'s method	0.01586	0.00434	0.00076	1044.14
	Our approach	0.00126	0.00310	0.00046	15.07
Dataset-1	Input	0.03910	0.01470	0.00179	0.00
	Park et al.'s method	0.02440	0.00597	0.00091	5209.64
	Our approach	0.02055	0.00327	0.00063	56.89
Dataset-2	Input	0.06835	0.02450	0.00248	0.00
	Park et al.'s method	0.03402	0.00895	0.00101	3649.97
	Our approach	0.02068	0.00232	0.00039	35.22
Dataset-3	Input	0.21265	0.03764	0.00491	0.00
	Park et al.'s method	0.09458	0.02145	0.00267	2744.31
	Our approach	0.07294	0.01207	0.00220	38.36

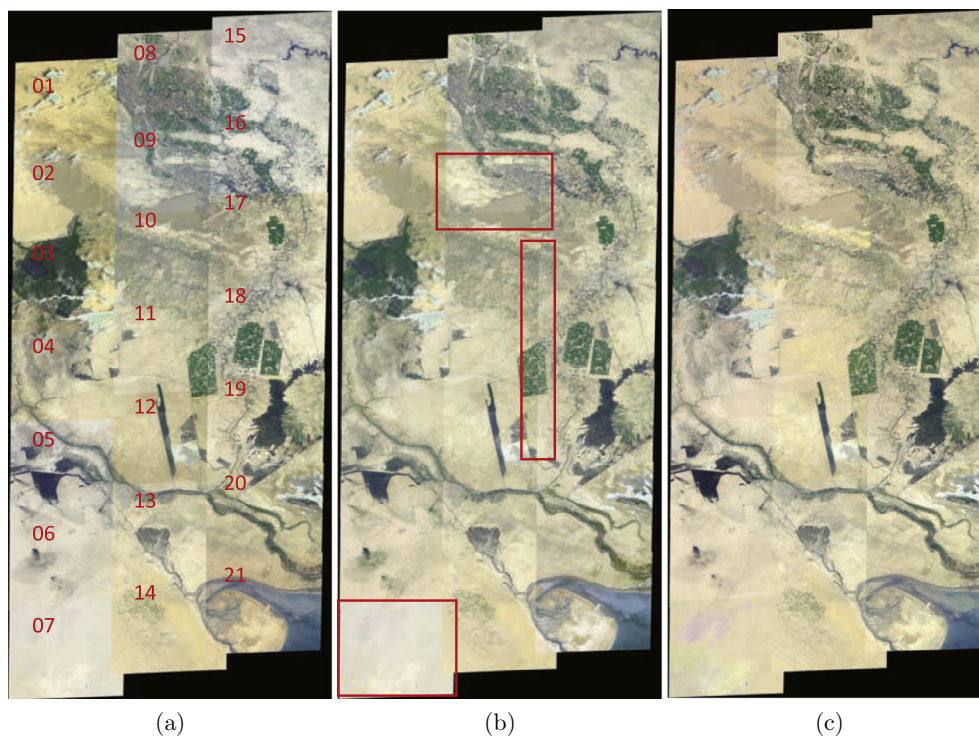


Fig. 12. The experimental results on three strips of multi-temporal images (Dataset-1): (a) original images labeled by its image index at the upper left corner; (b) the result corrected by Park et al.'s approach; (c) the result corrected by our approach with the auto-selected consistent reference image subset {13, 14, 19, 20, 21}.

in the first and second strips are more obvious, on the contrary, the color consistency optimization performed better between the third strip of images and the fourth strip of images. It is because the former has lower overlaps and feature matching didn't work so well, which resulted in a few feature matches. Although there is a great color difference among original images, the result corrected by our approach takes on such a good consistent appearance that the boundary lines between images are almost invisible. Also, we observe that the color consistency was kept well in intra-strip images. Table 1 shows the comparative statistics in mean histogram distances between all pairs of images for Park et al.'s approach and ours. It should be noted that the reference images were selected manually as {03, 04, 09, 10}, because setting the automatic consistent subset as reference would result in a little over-saturation effect on the image 04 (for example, the lake turns sort of green). This is the weakness of histogram mapping algorithm for color transfer, which is prone to possibly causing the

pixel saturation problem when the color difference among images is very large or the color transferring path is too long.

The last dataset contains 18 multi-temporal images acquired by Chinese GF-2 satellite, which present obvious tonal variations as shown in Fig. 14(a). The corresponding results of Park et al.'s approach and ours are illustrated in Fig. 14(b) and (c) respectively. Obviously, the whole visual feeling of Fig. 14(c) is more natural and consistent than that of Fig. 14(b), as can be observed that Fig. 14(b) improves very limitedly with respect to Fig. 14(a). This was contributed by two aspects: the one is that the correction model is more flexible to handle dramatic color disparity, and another is that the transfer of reference tone guarantees a natural appearance.

4.2.2. Efficient analysis

Our method was implemented with C++, while the code of Park et al.'s approach is provided in Matlab implementation. All the procedures were tested in a computer with an Intel Core i5-3350 at

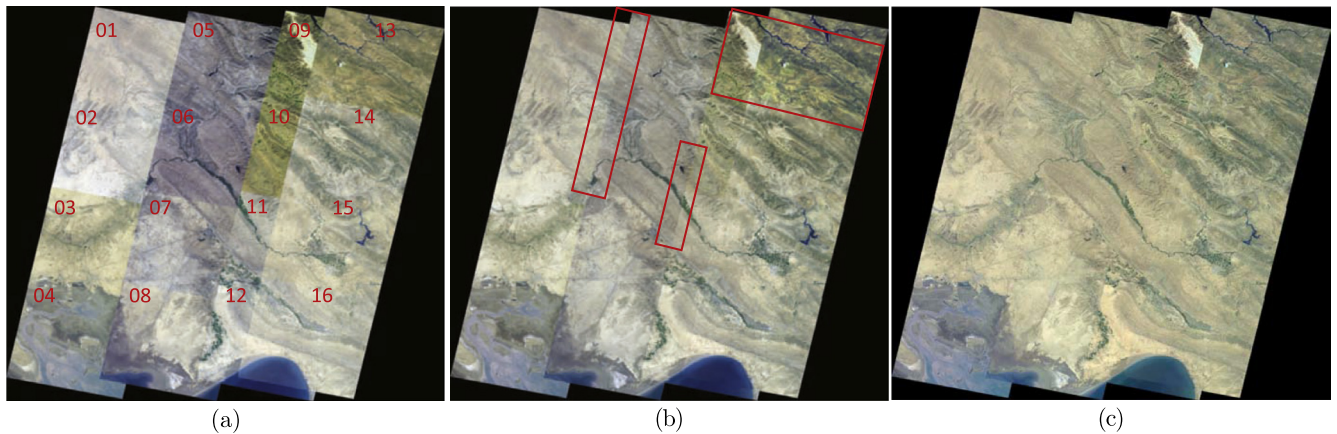


Fig. 13. The experimental results on four strips of multi-temporal images (Dataset-2): (a) original images with large and diverse color differences; (b) the result corrected by Park et al.'s approach; (c) the result corrected by our approach with the manually selected yellow-tone reference images {03, 04, 09, 10}. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

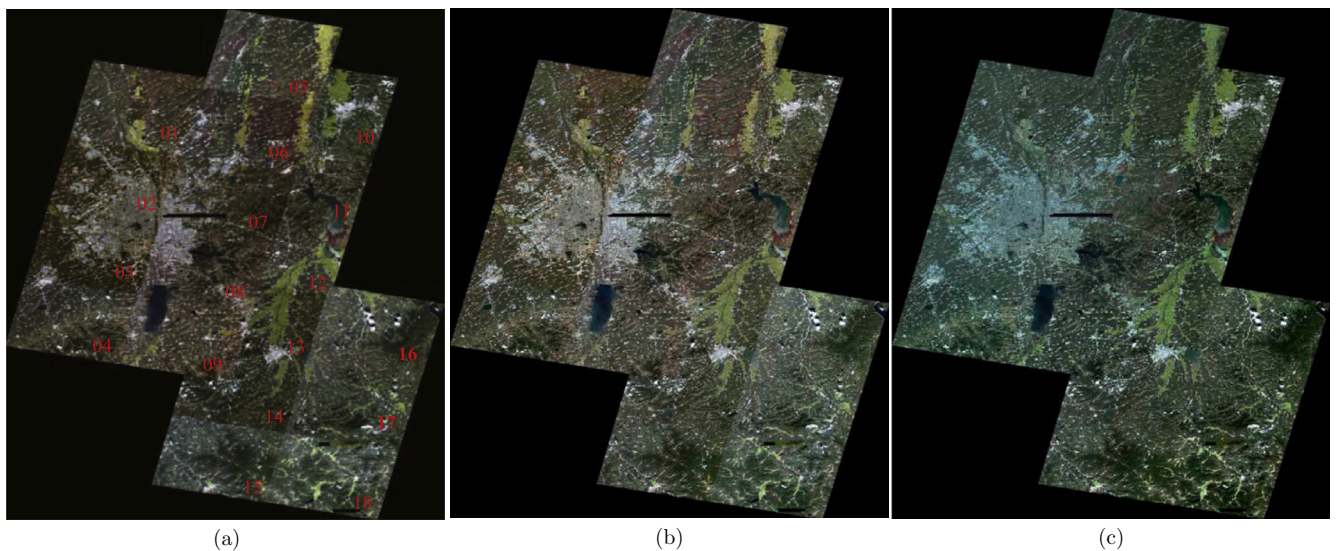


Fig. 14. The experimental results on multi-temporal images acquired by Chinese GF-2 satellite (Dataset-3): (a) original images with large and diverse color differences; (b) the result corrected by Park et al.'s approach; (c) the result corrected by our approach with the color consistent reference image subset {10, 16, 17, 18}. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.1 GHz with 16 GB RAM and the running times are depicted in Table 1. Because the publicly provided code of Park et al.'s approach was implemented with Matlab, the comparison of running time is meaningless and we prefer to analyze the algorithm complexity. Park et al. utilized a robust low-rank matrix factorization method to estimate model parameters, which just needs sparse feature matching. But, structure from motion (SfM) and inner iterative scheme are respectively used to obtain multi-image correspondence and matrix decomposition, which take the major computation and the computational cost towards time is typically high. In our method, color correspondences extraction takes the major computation, and the global optimization converges quickly with the provided initial solution.

In general, an image featuring or blending operation must be applied after color tonal correction to generate a mosaic as seamless as possible. To demonstrate the effectiveness and superiority of our method, we applied a multi-band blending method (Burt and Adelson, 1983) with an open-source software called “Enblend”¹

on the corrected images from the second real dataset. Enblend allows the user to change the number of levels for pyramid blending, based on which we can change the number of levels to analyze the color consistency performance on the finally mosaicked results. Each new level works on a linear scale twice as large as the previous one, which means the n -th level contains the image data at the 2^n -pixel scale. The image with the size of $width \times height$ pixels can not be deconstructed into a pyramid of more than $\log_2(\min(width, height))$ levels. Fewer levels emphasize local features and suppress global ones. For comparison, we applied the multi-band blending on three groups of images, consisting of the original second real dataset, and the corrected results on the second real dataset via Park et al.'s approach and ours. Limited by the image size, the number of levels only can be changed from 1 to 11 and the mosaicked results with 1, 6 and 11 levels are illustrated in Fig. 15. The mosaicked images in the first row, as shown Fig. 15(a)–(c), were generated by applying Enblend with 1 level, which is equivalent to the simple superposition with the detected seamlines in practice. The stitching artifacts caused by color inconsistencies can be obviously observed in Fig. 15(a)–(b) but it almost disappeared in our result as shown in Fig. 15(c). With the increase of the number of levels, the stitching

¹ Available at <http://enblend.sourceforge.net/>.

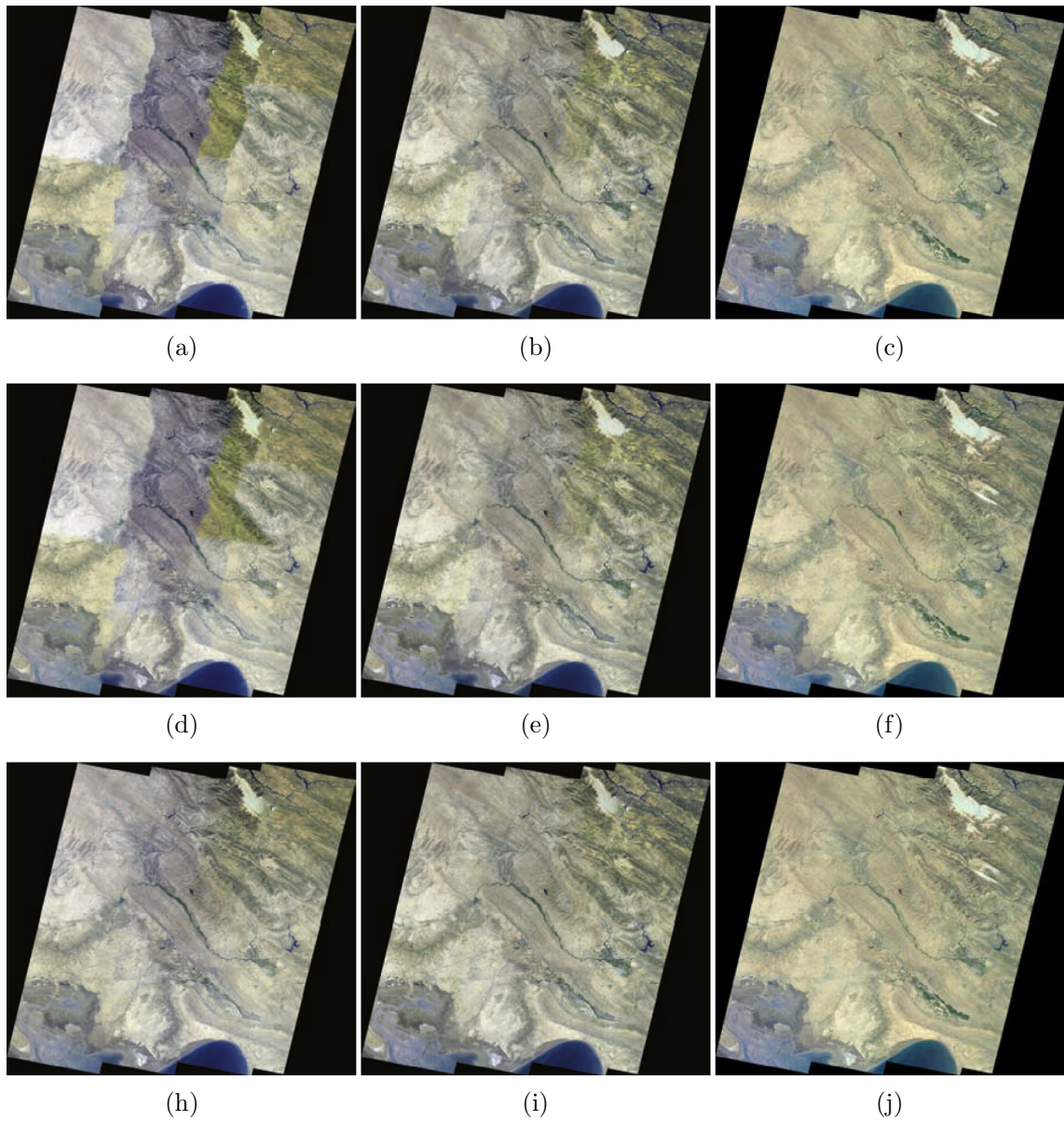


Fig. 15. Visual comparison with respect to different numbers of levels used in “Enblend”: (a), (d) and (h) the results of original images after blending; (b), (e) and (i) Park et al.’s results after blending; (c), (f) and (j) the results of our approach after blending. From top to bottom, the comparative results correspond to one, six and eleven levels, respectively.

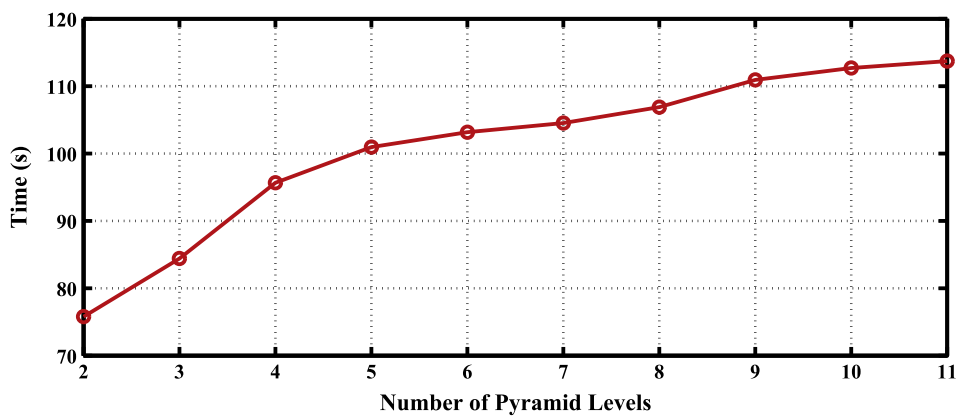


Fig. 16. The computational times of applying “Enblend” with different numbers of pyramid levels on the second real multi-temporal image dataset shown in Fig. 15(a).

artifacts were gradually concealed. Even with the maximum number (i.e., 11) of levels, the stitching artifacts can be observed indistinctly on the results of the original images and the corrected ones by Park et al.'s approach as shown in Fig. 15(h) and (i), respectively. This experiment demonstrates that the color consistency correction is necessary for image mosaicking when there exist large color differences among the images. In addition, it sufficiently demonstrates that our approach outperformed Park et al.'s.

The computational costs of applying Enblend with different numbers of pyramid levels for image blending are shown in Fig. 16. These costs don't cover the computation of seamline searching. The computation of applying Enblend with 1 level was not taken into account, because it is equivalent to the simple superposition with the detected seamlines in practice. From the experimental results of different numbers of pyramid levels, we can observe that the computational time ascends along with the

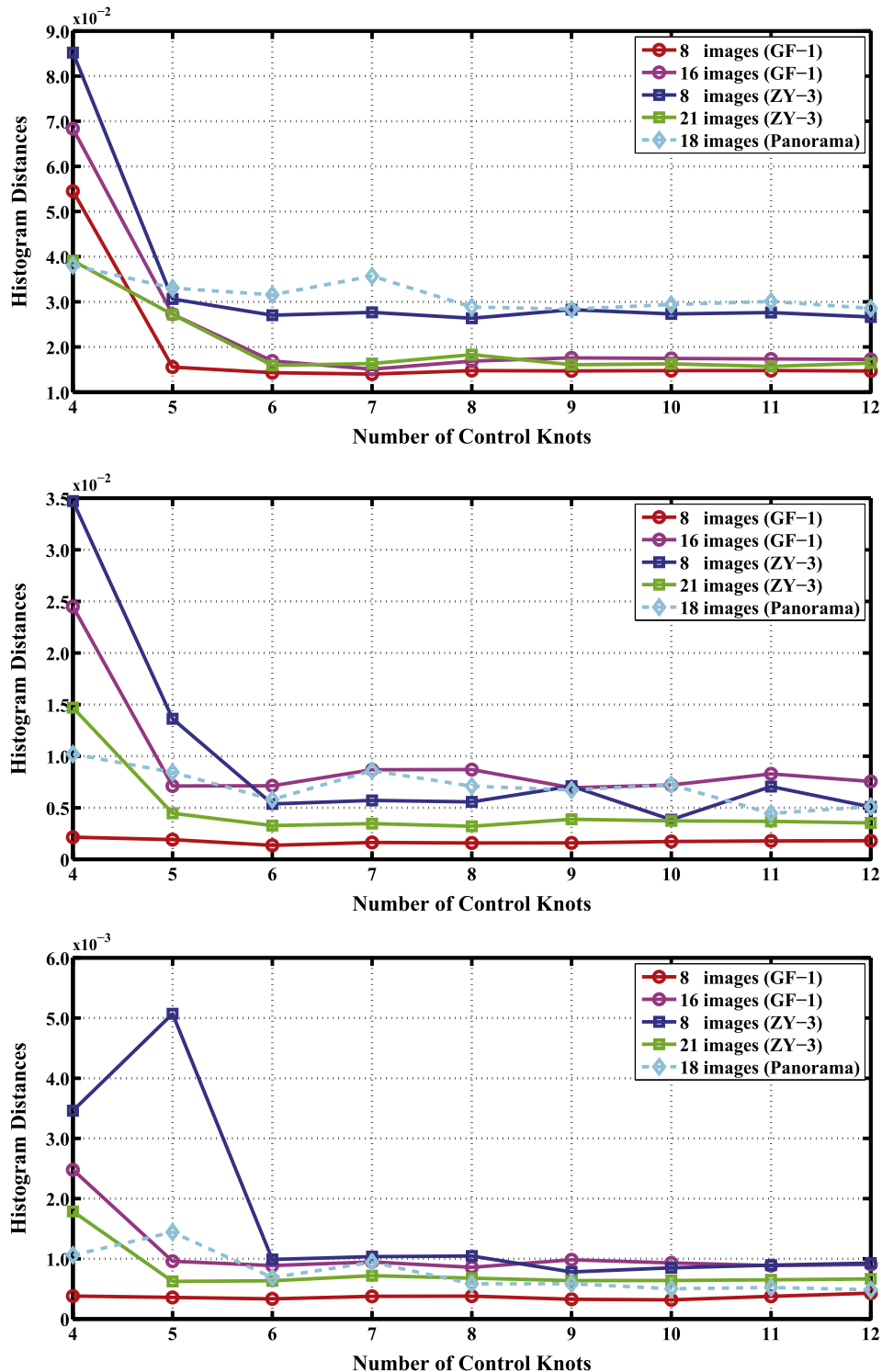


Fig. 17. Comparative results on histogram distances of different image sets with respect to the numbers of control knot. From top to bottom, the comparative results correspond to the α , β , and γ channels, respectively. The solid and dotted lines correspond to remote sensing image sets and the panoramic image set, respectively.

increase of the number of pyramid levels. In order to achieve a good color consistency of an image mosaic, the required number of pyramid levels corresponding to our approach is much less than the number of Park et al.'s approach, which results in the less computation cost. These experimental results demonstrate the superiority of our approach over Park et al.'s in the aspect of computational effectiveness.

4.3. Analysis on the number of control knots

We selected several datasets to analyze the number of evenly-distributed control knots that parameterize the quadratic spline curve. To keep the diversity of testing data, both panoramic images and remote sensing ones were used to analyze the effect of the number of control knots. Considering more obvious color differences existed in remote sensing images and the just different luminances existed in the most panoramic images, we chose a set of 18 panoramic images as shown in Fig. 8 and four sets of remote sensing ones including multi-temporal and multi-resolution images with two or three strips consisting of 8 and 16 ones captured by the Chinese satellite GF-1, and 8 and 21 ones captured by the Chinese satellite ZY-3.

When the number of control knots decreases for the same dataset, the corresponding constraint on the quadratic spline curve will be loosened, which possibly results in the lower optimization accuracy. In theory, we can choose a greater number of control knots to increase the accuracy, but the corresponding optimization becomes more complex. The comparative analysis on five image sets reported in Fig. 17 shows that the histogram distances of the overlapping image regions for each channel (corresponding to α and β , respectively) decrease quickly at the beginning and keep relatively stable when the number of control knots exceeds 6. Taking both accuracy and complexity into consideration, we used 6 control knots in this paper.

5. Conclusion

In this paper, we proposed a novel approach to improve the color consistency of sequential images for mosaicking. Firstly, the specially proposed histogram extreme points matching strategy increases the reliability and precision of intensity correspondence greatly. To achieve a satisfying consistent appearance of the images, we proposed to provide an initial solution, learned from the major color style of the original images, to the global optimization as a guidance. The energy function of the global optimization was designed to minimize the total color difference under the constraint of not going far away from the initial solution. Experimental results on several representative synthetic and real image datasets sufficiently demonstrate the proposed method is superior to the state-of-the-art algorithm through qualitative analysis and quantitative estimation. However, there also exist some deficient aspects to be further improved in our approach. One is the strategy of selecting the reference image subset, which won't select two or multiple subsets disconnected with each other, even though they are in a very consistent tone. Besides, when the color difference among images is very large or the color transferring path is too long, the weakness of histogram mapping algorithm for color transfer is prone to causing pixel saturation problem. A more effective solution to avoid pixel saturation should be considered in such algorithm based on the mapping curve optimization. All of these problems will be deeply investigated in our future work.

Acknowledgment

This work was partially supported by the National Natural Science Foundation of China (Project No. 41571436), the Hubei

Province Science and Technology Support Program, China (Project No. 2015BAA027), the National Natural Science Foundation of China under Grant 91438203, LIESMARS Special Research Funding, and the South Wisdom Valley Innovative Research Team Program.

References

- Allene, C., Pons, J., Keriven, R., 2008. Seamless image-based texture atlases using multi-band blending. In: IEEE International Conference on Pattern Recognition (ICPR).
- Baudisch, P., Gutwin, C., 2004. Multiblending: displaying overlapping windows simultaneously without the drawbacks of alpha blending. In: Sigchi Conference on Human Factors in Computing Systems.
- Brown, M., Lowe, D.G., 2007. Automatic panoramic image stitching using invariant features. *Int. J. Comput. Vision* 74, 59–73.
- Burt, P., Adelson, E.H., 1983. A multiresolution spline with application to image mosaics. *ACM Trans. Graph.* 2 (4), 217–236.
- Canty, M.J., Nielsen, A.A., 2008. Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. *Remote Sens. Environ.* 112 (3), 1025–1036.
- Canty, M.J., Nielsen, A.A., Schmidt, M., 2004. Automatic radiometric normalization of multitemporal satellite imagery. *Remote Sens. Environ.* 91 (3), 441–451.
- Frigo, O., Sabater, N., Demoulin, V., Hellier, P., 2011. Optimal transportation for example-guided color transfer. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Vol. 9005 of Computer Vision – ACCV 2014, pp. 655–670.
- Goldman, D.B., 2010. Vignette and exposure calibration and compensation. *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (12), 2276–2288.
- HaCohen, Y., Shechtman, E., Goldman, D.B., Lischinski, D., 2011. Non-rigid dense correspondence with applications for image enhancement. *ACM Trans. Graph. (TOG)* 30 (4), 70:1–70:10.
- HaCohen, Y., Shechtman, E., Goldman, D.B., Lischinski, D., 2013. Optimizing color consistency in photo collections. *ACM Trans. Graph.* 32 (4), 38:1–38:10.
- Hu, W., Xie, N., Hu, R., Ling, H., Chen, Q., Yan, S., Maybank, S., 2014. Bin ratio-based histogram distances and their application to image classification. *IEEE Trans. Pattern Anal. Mach. Intell.* 36 (12), 2338–2352.
- Huang, T.-W., Chen, H.-T., 2009. Landmark-based sparse color representations for color transfer. In: IEEE International Conference on Computer Vision (ICCV).
- Hwang, Y., Lee, J.-Y., Kweon, I.S., Kim, S.J., 2014. Color transfer using probabilistic moving least squares. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 3342–3349.
- Ilie, A., Welch, G., 2005. Ensuring color consistency across multiple cameras. In: IEEE International Conference on Computer Vision (ICCV).
- Kerschner, M., 2001. Seamline detection in colour orthoimage mosaicking by use of twin snakes. *ISPRS J. Photogramm. Remote Sens.* 56, 53–64.
- Kim, S.J., Pollefeys, M., 2008. Robust radiometric calibration and vignetting correction. *IEEE Trans. Pattern Anal. Mach. Intell.* 30 (4), 562–576.
- Levin, A., Zomet, A., Peleg, S., Weiss, Y., 2004. Seamless image stitching in the gradient domain. In: European Conference on Computer Vision (ECCV). Springer.
- Li, X., Hui, N., Shen, H., Fu, Y., Zhang, L., 2015. A robust mosaicking procedure for high spatial resolution remote sensing images. *ISPRS J. Photogramm. Remote Sens.* 109, 108–125.
- López, D.H., García, B.F., Piqueras, J.C., Alcázar, G.V., 2011. An approach to the radiometric aerotriangulation of photogrammetric images. *ISPRS J. Photogramm. Remote Sens.* 66 (6), 883–893.
- Moulon, P., Dussat, B., Monasse, P., 2013. Global multiple-view color consistency. In: International Conference on Visual Media Production (CVMP).
- Naim, M.-J.-N.M., Isa, N.A.M., 2012. Pixel distribution shifting color correction for digital color images. *Appl. Soft Comput.* 12 (9), 2948–2962.
- Oakley, J.P., Bu, H., 2007. Correction of simple contrast loss in color images. *IEEE Trans. Image Process.* 16 (2), 511–522.
- 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.
- Papadakis, N., Provenzi, E., Caselles, V., 2011. A variational model for histogram transfer of color images. *IEEE Trans. Image Process.* 20 (6), 1682–1695.
- Park, J., Tai, Y.-W., Sinha, S.N., Kweon, I.S., 2016. Efficient and robust color consistency for community photo collections. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 430–438.
- Perez, P., Gangnet, M., Blake, A., 2003. Poisson image editing. *ACM Trans. Graph. (TOG)* 22 (3), 313–318.
- Pérez, P., Gangnet, M., Blake, A., 2003. Poisson image editing. In: ACM Transactions on Graphics.
- Pitié, F., Kokaram, A.C., Dahyot, R., 2005. N-dimensional probability density function transfer and its application to color transfer. In: IEEE International Conference on Computer Vision (ICCV), pp. 1434–1439.
- Pitié, F., Kokaram, A.C., Dahyot, R., 2007. Automated colour grading using colour distribution transfer. *Comput. Vis. Image Underst.* 107 (1), 123–137.
- Popovic, V., Afshari, H., Schmid, A., Leblebici, Y., 2013. Real-time implementation of gaussian image blending in a spherical light field camera. In: IEEE International Conference on Industrial Technology.
- Prados, R., García, R., Neumann, L., 2013. Image blending techniques and their application in underwater mosaicking. *Springerbriefs in Computer Science*.

- Prados, R., Garcia, R., Neumann, L., 2014. *Image Blending Techniques and Their Application in Underwater Mosaicing*. Springer.
- Reinhard, E., Ashikhmin, M., Gooch, B., Shirley, P., 2001. Color transfer between images. *IEEE Comput. Graph. Appl.* 21 (5), 34–41.
- Rizzi, A., Gatta, C., Marini, D., 2003. A new algorithm for unsupervised global and local color correction. *Pattern Recogn. Lett.* 24 (11), 1663–1677.
- Shao, F., Jiang, G.-Y., Yu, M., Ho, Y.-S., 2010. Fast color correction for multi-view video by modeling spatio-temporal variation. *J. Vis. Commun. Image Represent.* 21 (5-6), 392–403.
- Snavely, N., Seitz, S.M., Szeliski, R., 2006. Photo tourism: exploring photo collections in 3d. *ACM Trans. Graph.* 25, 835–846.
- Szeliski, R., Uyttendaele, M., Steedly, D., 2011. Fast Poisson blending using multi-splines. In: *IEEE International Conference on Computational Photography*.
- Tai, Y.-W., Jia, J., Tang, C.-K., 2005. Local color transfer via probabilistic segmentation by Expectation-Maximization. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 747–754.
- Tehrani, M.P., Ishikawa, A., Sakazawa, S., Koike, A., 2010. Iterative colour correction of multicamera systems using corresponding feature points. *J. Visual Comm. Image Represent.* 21 (5-6), 377–391.
- Tian, G.Y., Gledhill, D., Taylor, D., Clarke, D., 2002. Colour correction for panoramic imaging. In: *Proceedings of the International Conference on Information Visualisation*.
- Vallet, B., Lelégard, L., 2013. Partial iterates for symmetrizing non-parametric color correction. *ISPRS J. Photogramm. Remote Sens.* 82 (8), 93–101.
- Wang, M., Pan, J., Chen, S., Li, H., 2005. A method of removing the uneven illumination phenomenon for optical remote sensing image. In: *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3243–3246.
- Xiang, Y., Zou, B., Li, H., 2009. Selective color transfer with multi-source images. *Pattern Recogn. Lett.* 30 (7), 682–689.
- Xiong, Y., Pulli, K., 2009. Mask based image blending and its applications on mobile devices. In: *Sixth International Symposium on Multispectral Image Processing and Pattern Recognition*.
- Xiong, Y., Pulli, K., 2010. Fast panorama stitching for high-quality panoramic images on mobile phones. *IEEE Trans. Consumer Electron.* 56, 298–306.
- Xu, W., Mulligan, J., 2010. Performance evaluation of color correction approaches for automatic multi-view image and video stitching. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 263–270.
- Yoo, J.-D., Park, M.-K., Cho, J.-H., Lee, K.H., 2013. Local color transfer between images using dominant colors. *J. Electron. Imaging* 22 (3), 033003.