Color Consistency Correction Based on Remapping Optimization for Image Stitching

Menghan Xia Jian Yao* Renping Xie Mi Zhang
School of Remote Sensing & Information Engineering, Wuhan University, Wuhan, Hubei, P.R. China
*Email: jian.yao@whu.edu.cn Web: http://cvrs.whu.edu.cn

Jinsheng Xiao
Electronic Information School, Wuhan University, Wuhan, Hubei, P.R. China

Abstract

Color consistency correction is a challenging problem in image stitching, because it matters several factors, including tone, contrast and fidelity, to present a natural appearance. In this paper, we propose an effective color correction method which is feasible to optimize the color consistency across images and guarantee the imaging quality of individual image meanwhile. Our method first apply well-directed alteration detection algorithms to find coherent-content regions in inter-image overlaps where reliable color correspondences are extracted. Then, we parameterize the color remapping curve as transform model, and express the constraints of color consistency, contrast and gradient in an uniform energy function. It can be formulated as a convex quadratic programming problem which provides the global optimal solution efficiently. Our method has a good performance in color consistency and suffers no pixel saturation or tonal dimming. Experimental results of representative datasets demonstrate the superiority of our method over state-of-the-art algorithms.

1. Introduction

As the growing popularity of image capture equipments and photo sharing, we now are in an image world where the image data can be obtained easily by using SLR cameras, smart phones, or downloading from the social networks and public data platforms[1]. This exciting data availability enables many visual applications, such as image recognition [7], panoramic imaging [1, 11, 27], image rendering [31, 32] and virtual navigation [19, 9]. In many cases, images of the same scene might show noticeable tonal inconsistence because of different atmosphere illumination, exposure time and camera response function. Such photometric difference could particularly influence the visual effects of multiple images based rendering mission [29, 18]. This paper focus on studying the color correction problem for multiview image stitching, as exemplified in Figure 1.

Figure 1: Panorama composited with source images taken by cameras with different imaging settings. Color inconsistence is still noticeable in the scene even processed by seamline selection and multi-band blending in *Enblend*[2].

Within the technique pipeline of image stitching, color correction is a critical step in presenting the composited image with an natural and consistent tone. Apart from removing the overall tonal disparity, it also facilitates the following seamline selection [10] and blending [22]. However, obtaining a satisfactory color correction result is a non-trivial task: For one thing, a rigorous correction model that can genuinely express the inter-image color transforming relations remains an open issue; For another, the original distribution of pixel value indicates semantic information (object structure, contrast, saturation, etc), which might might be degraded or destroyed when adjusting pixels for the least color disparity. This is a solved problem for two image involved color transfer techniques that allows complex operations to enhance process quality, however it is hard to be extended to multiple images scheme uniformly. Cascading manipulation strategy [14, 26] subjects to linearly amplified accumulation error, making the global consistency inacces-

[1]USGS (satellite image): http://glovis.usgs.gov/

[2]Available at: http://enblend.sourceforge.net/

sible. As for multiview color correction, existing methods mainly employed simple correction models or neglected detail preservation and contrast consideration in the global optimization for consistency [29, 18]. Thus, a robust and effective multiview color correction algorithm remains to be investigated further.

As to the problems stated above, this paper presents an effective method integrating detail preservation, global contrast and color consistency into an uniform optimization framework. It can guarantee the optimality of the processing result. Like regular approaches, we leverage the color statistics of overlap regions to obtain the color correspondences. Specially, to improve the accuracy and robustness of this, the classic Iteratively Reweighted Multivariate Alteration Detection (IR-MAD) [13] is used to detect and remove the disturbance of altered content in overlaps. Inspired by the flexible model used in [5], we parameterize the quadratic spline curve as transform function, which can express different constraints in the optimization function effectively. Besides, the optimization problem can be solved efficiently by virtue of convex quadratic programming. Through comparing to the latest methods, our approach illustrated better performance in color consistency and dynamic range optimization.

The rest of the paper is structured as follows. In Section 2 an overview of related work is given. In Section 3 the scope of the problem is formally defined and in Section 4 the proposed color correction approach is detailed. In Section 5 experimental results are evaluated. Finally, conclusions and future work are presented in Section 6.

2. Related Works

Regarding color processing, there are two relevant techniques: two image based color transfer [4, 8], and multiview color correction [1, 6, 15]. Existing methods are reviewed below respectively.

2.1. Color Transfer

The concept of color transfer was first proposed by Reinhard *et al.* [4], which aims at propagating one image's color characteristic to another. From this baseline approach, many other following works were proposed. Before 2010, they mainly focus on solving the decorrelation between color channels [16] and content-based elaborating color transferring [23], which were well summarized in [30]. Since then, more emphasis was laid on grain-free, detail preservation and artifacts suppression. To heighten the detail performance, Xiao *et al.* [28] proposed a gradient-preserving model to convert the transfer processing to an optimization balancing the color distribution and the detail performance. Different from a stepwise strategy, Su *et al.* [20] proposed to perform color mapping and detail boosting separately through gradient-aware decomposition, to obtain a grain-free and detail preserved result. Based on this idea, they developed more complete framework to suppress corruptive artifacts [21].

To avoid color distortion, Nguyen *et al.* [12] applied Xiao *et al.*'s algorithm [28] in luminance channel, followed by the color gamut alignment via rotating around luminance axis. It was effective in preventing color distortion but lacking in color fidelity because of its simple color transform model. To make a more accurate color mapping, Hwang *et al.* [8] proposed to correct each pixel's color with an independent affine model, which is the solution of probabilistic moving least square based on feature color correspondences. Besides, some works presented high-level color transfer application based on semantic segmentation or content recognition [25, 3]. Different with these traditional methods, learning-based color transfer methods were attempted to train out the proper color mapping relationship [24, 2].

2.2. Multiview Color Correction

Here, multiview color correction refers to correct the color of three or more images for consistency in a global way, which excludes applying color transfer across image sequence or network to realize color consistency. Brown *et al.* [1] first proposed to performed the gain compensation for multiple images in the way of global optimization, which has been applied in panorama software *Autostitch*. Under the linear optimization framework, Xiong *et al.* [29] extended this method by employing gamma model in luminance channel and linear model in chromatic channels. Qian *et al.* [17] presented a manifold method to remove color inconsistence, which made full use of both the correspondences in overlap regions and the intrinsic color structures. However, its requirement on accurate geometric alignment limited its application range.

In 3D modeling, Shen *et al.* [18] exploited linear function as color correction model over color histogram to generate consistent textures. It is efficient but unable to process great color discrepancy. In photo editing applications, Hacohen *et al.* [6] proposed to model the remapping curve directly, which were optimized for consistent appearance of photos using color correspondences obtained from non-rigid dense matching [5]. This color model is flexible enough to correct even large tonal disparity but its dependence on dense matching makes it computationally expensive. To address it, Park *et al.* [15] presented a robust low-rank matrix factorization method to estimate its correction parameters, which just need sparse feature matching. But, its color correction ability is not as high.

3. Problem Formulation

First of all, our color correction method aims to be applied on the sequential images whose aligning models have

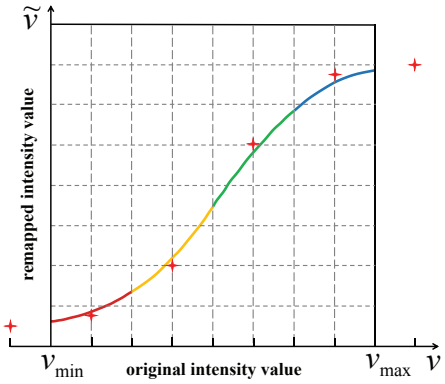


Figure 2: Piecewise quadratic spline: the color mapping function used in our method. Leveraging the fixed horizontal distribution of red anchor knots, the curve just covers the domain of the original intensity values $[v_{\min}, v_{\max}]$.

been estimated in advance. Thus, the adjacent relationships and overlap regions are used as given information in our approach. Second, our color correction algorithm runs on the assumption that all the color inconsistence among images is caused by different imaging conditions which affect each image as a whole.

Seeking for the optimal consistency, our approach expresses all the quality requirements in the form of constraint on model parameters, which are then solved in a global optimization. To realize this, we define the transformation model as three monotonically increasing mapping curves (one per channel), each of which is formulated as a piecewise quadratic spline with m control knots ($m = 6$ as default). As illustrated in Figure 2 , these red knots $\{(\nu_k, \tilde{\nu}_k)\}_{k=1}^{m}$ are half free on the coordinate plane, where $\{\nu_k\}_{k=1}^{m}$ are fixed evenly on the horizontal axis to control the mapping curve effectively, while $\{\tilde{\nu}_k\}_{k=1}^{m}$ are free to determine the shape of the mapping curve as the actual model parameters. Thus, the color mapping function for image $\mathbf{I}_i$ can be parameterized as:

$$F_i = \arg\{(\tilde{\nu}_1^i, \tilde{\nu}_2^i, ..., \tilde{\nu}_m^i)_c\}_{c=1}^{3}, \tag{1}$$

where c is the index number of each channel. Particularity, having the curve cover the intensity domain of each original image can save the troublesome extrapolation for non-overlap regions. Besides, the detailed definition of our optimization function are described in Section 4.2.

4. Color Correction

Our color correction method consists of two steps: color correspondence extraction and model parameters optimization. Particularly, the color correction is performed in $YCbCr$ space, because its luminance channel and chromatic channels are separated and each channel has specific upper and lower bounds, which facilitate the quality-aware color consistency optimization.

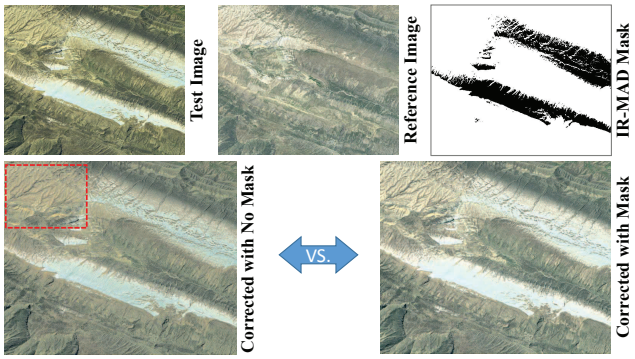


Figure 3: Visual comparison between the corrected result with alteration filtering mask applied and that with none mask applied when alteration exists.

4.1. Reliable Color Correspondence

Based on the adjacent relations, we extract color correspondences in the overlaps of each image pair. For efficiency problem, we adopt the statistical measures of color histogram, instead of matching image color by pixels. That's to say, we take the same quantiles (color value pairs of the same frequency) in the cumulative color histograms of the shared contents as correspondences. In general case, the overlap regions can be regarded as the shared contents, even if a bit of misalignment exists. However, if obviously altered objects exist there, their pixels as outliers should be excluded from the overlaps in advance of histogram counting. To do so, we utilize a famous alteration detection algorithm IR-MAD [13] in overlap regions to generate filtering masks. However, it is computationally expensive to run IR-MAD on the images of their original size, especially for high-resolution remote sensing images. Thus, we downsample each overlap region into the size of 250K pixels at first and then apply the IR-MAD on them to get a coarse mask. Here is an example that applying alteration filtering mask improves the accuracy of color correction result when alteration exists in Figure 3.

4.2. Remapping Model Optimization

With color correspondences of image pairs, energy function can be designed by using the parameters that remain to be optimized. Given a group of images $\{\mathbf{I}_i\}_{i=1}^{n}$, we attempt to seek a group of color transformations $\{F_i\}_{i=1}^{n}$ which maintains a good balance between three quality goals:

- *color:* pixels depicting the same content have the same color across images;
- *gradient:* original structural details remain on the transformed images;
- *contrast:* all the transformed images have a reasonable wide dynamic range;

As stated in Section 3, all these quality requirements should be expressed as the constraints on parametric model. In our method, the global optimization is conducted independently in each channel. Without losing generality, the transformation function F_i of $\mathbf{I}_i$ in certain channel is denoted as $f_i = \arg(\tilde{\nu}_1^i, \tilde{\nu}_2^i, ..., \tilde{\nu}_m^i)$ for simplicity. So, in any channel, the energy function can be formulated as:

$$E = \sum_{\mathbf{I}_i \cap \mathbf{I}_j \neq \emptyset} w_{ij} E_{\text{data}}(f_i, f_j) + \lambda \sum_{i=1}^{n} E_{\text{regulation}}(f_i)$$
$$\text{subject to: } C_{\text{rigid}}(f_i), \forall i \in [1, n], \quad (2)$$

where w_{ij} is the weight proportional to the area of the overlap region between $\mathbf{I}_i$ and $\mathbf{I}_j$ ($\sum w_{ij} = 1$), and λ serves to balance the data term and the regulation term. In our experiments, $\lambda = \xi \frac{M}{m}$ is used, where M denotes the amount of color correspondences extracted in each overlap ($M = 16$ as default) and $\xi \in [0.5, 5]$ is recommended. Actually, our algorithm is insensitive to λ since this two terms are non-contrary. The color consistency across images is expressed by data term that penalizes the deviation between remapped corresponding color values:

$$E_{\text{data}}(f_i, f_j) = \sum\nolimits_{k=1}^{M} \left\| f_i(v_k^i) - f_j(v_k^j) \right\|^2, \quad (3)$$

where $\{v_k^i, v_k^j\}_{k=1}^M$ are color correspondences between $\mathbf{I}_i$ and $\mathbf{I}_j$. As an unary term, $E_{regulation}(f_i)$ enforces certain constraints on the color transformations softly, including regularizing the parameters and stretching dynamic range.

$$E_{\text{regulation}}(f_i) = \sum\nolimits_{k=1}^{m-1} \left\| f_i(\dot{v}_k^i) - \dot{v}_k^i \right\|^2 - \eta \| f_i(v_{0.05}^i) - f_i(v_{0.95}^i) \|^2, \quad (4)$$

where $\{\dot{v}_k^i\}_{k=1}^{m-1}$ denote the horizontal coordinates of joint points joining different local curve segments (marked in different colors in Figure 2). Exactly, we have $\dot{v}_k^i = \frac{\nu_k^i + \nu_{k+1}^i}{2}, k = \{1, 2, ..., m-1\}$. The importance of this term lies in two aspects: (i) keeping transformed images close to their original appearance slightly as default solution; (ii) all the model parameters are optimized as a whole even if no color correspondences falls into the scope of some anchors (parameters). In addition, the latter negative term prevents the dynamic range to narrow down where v_α^i depicts the α-percentile of the CDF of $\mathbf{I}_i$. This term is very necessary to avoid an easy-appearing optimizing result that corrected images present a consistent but dimming tone, because remapping intensity into a lower range conforms to the minimal energy principle of E_{color}. Particularly, $\eta = 5$ and $\eta = 0$ are used in luminance channel and chromatic channels respectively. Namely, we only impose dynamic range constraint in luminance.

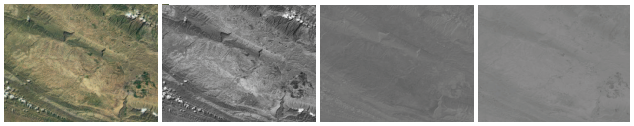


Figure 4: Channel information display of a color image in $YCbCr$ space. The major gradient information is contained in luminance channel Y.

As a qualified color mapping curve, the quadratic spline should meet two basic requirements: increasing monotonicity and mapping domain lying within gamut. They are guaranteed through the following constraint term:

$$C_{\text{rigid}}(f_i) : \begin{cases} \tau_b \leq f_i'(v_k^i) \leq \tau_u, \ \forall v_k^i \in [v_{\min}, v_{\max}], \\ v_{\text{start}} \leq f_i(v_{0.01}^i), \ f_i(v_{0.99}^i) \leq v_{\text{end}}, \end{cases} \quad (5)$$

where τ_b and τ_u defines the bottom and upper boundary of the mapping curve's slope domain, and $[v_{\text{start}}, v_{\text{end}}]$ is the gamut range of related channel. Actually, the definition of slop domain, especially its bottom boundary, is a trade-off problem between detail preservation and color consistency, since a higher value of bottom boundary (such as $\tau_b \geq 1.0$) can prevent gradient detail loss from intensity levels merging, while it also inevitably restricts the freedom of parameters to achieve a better color consistency. In our method, we set the slop domain $[0.3, 5]$ for chromatic channels Cb and Cr , and $[0.5, 5]$ for luminance channel Y where most of the gradient information is contained, as the example shown in Figure 4.

4.3. Implementation Details

Our correction function, as a piecewise interpolation model, can not be expressed explicitly. We now describe how to express the function mapping f_i and partial derivative f_i' with the actual model parameters $\{\tilde{\nu}_1^i, \tilde{\nu}_2^i, ..., \tilde{\nu}_m^i\}$.

Given a color value v_k^i, supposing it falls in the control scope of knots $\{(\nu_p^i, \tilde{\nu}_p^i)\}_p^{p+2}$, then its remapped value $\tilde{v}_k^i$ can be obtained via the following quadratic spline interpolation equation:

$$\begin{cases} v_k^i = \frac{1}{2}[(1-2t+t^2)\nu_p^i + (1+2t-2t^2)\nu_{p+1}^i + t^2\nu_{p+2}^i], \\ \tilde{v}_k^i = \frac{1}{2}[(1-2t+t^2)\tilde{\nu}_p^i + (1+2t-2t^2)\tilde{\nu}_{p+1}^i + t^2\tilde{\nu}_{p+2}^i], \end{cases} \quad (6)$$

where the only unknowns are interpolation coefficient $t \in [0, 1]$ and $\tilde{v}_k^i$, so $\tilde{v}_k^i$ can be expressed by model parameters $\{\tilde{\nu}_p^i\}_p^{p+2}$, which is the essence of f_i. In fact, we tend to solve the interpolation coefficients of all color correspondences before the global optimization, then the function mapping equals to a linear interpolation.

According to Eq. (6), we can express the slope value of

remapping curve at v_k^i as:

$$f_i'(v_k^i) = \frac{\partial \tilde{v}^i/\partial t}{\partial v^i/\partial t} = \frac{(\tilde{\nu}_p^i - 2\tilde{\nu}_{p+1}^i + \tilde{\nu}_{p+2}^i)t + \tilde{\nu}_{p+1}^i - \tilde{\nu}_p^i}{(\nu_p^i - 2\nu_{p+1}^i + \nu_{p+2}^i)t + \nu_{p+1}^i - \nu_p^i}, \tag{7}$$

where $\nu_p^i - 2\nu_{p+1}^i + \nu_{p+2}^i = 0$ since $\{\nu_p^i\}_{i=1}^m$ are constants distributing on the horizontal axis evenly (as shown in Figure 2). So, $f_i'(v_k^i)$ equals to a linear function with $t \in [0, 1]$ as the free variable. Then, the domain constraint $f_i' \in [\tau_b, \tau_u]$ can be expressed as linear inequalities:

$$\begin{cases} \tau_b \le \dfrac{\tilde{\nu}_{p+1}^i - \tilde{\nu}_p^i}{\nu_{p+1}^i - \nu_p^i} \le \tau_u, \ t = 0, \\ \tau_b \le \dfrac{\tilde{\nu}_{p+2}^i - \tilde{\nu}_{p+1}^i}{\nu_{p+1}^i - \nu_p^i} \le \tau_u, \ t = 1, \end{cases} \tag{8}$$

So far, we can substitute Eq. (6) and Eq. (8) into Eq. (3), Eq. (4) and Eq. (5). Then, the energy function Eq. (2) turns a quadratic polynomial, which can be transformed to the standard form of constrained quadratic programming, then is minimized by *convex quadratic programming*[3] efficiently.

5. Experimental Results

We compare the proposed approach against two latest methods in correcting color consistency of multiple images. The one is linear model based Shen *et al.*'s method [18], the other is albedo reflect model based Park *et al.*'s method [15]. As two typical applications of image stitching, ground panorama and remote sensing image mosaicking are used to evaluate the performance of color correction approaches. We use the default values of parameters described in Section 4.2 for all the experiments.

Generally, making a visually pleasing consistent result is the primary goal of color correction for image stitching. Thus, visual effects serve as the major evaluation criterion of algorithms. As supplement, quantitative evaluations on color Discrepancy (CD) and gradient loss (GL) are also conducted, as the metric equations show:

$$\begin{cases} \mathrm{CD} = \displaystyle\sum_{\mathbf{I}_i \cap \mathbf{I}_j \neq \emptyset} \bar{w}_{ij} \frac{\triangle \mathrm{H}(\hat{\mathbf{I}}_{ij}, \hat{\mathbf{I}}_{ji})}{N_{\mathrm{bin}}}, \\ \mathrm{GL} = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \frac{\triangle \mathrm{Go}(\mathbf{I}_i, \hat{\mathbf{I}}_i)}{N_{\mathrm{pix}}} \end{cases}, \tag{9}$$

where $\hat{\mathbf{I}}_i$ depicts the corrected image of source image $\mathbf{I}_i$, and $\hat{\mathbf{I}}_{ij}$ denotes the overlap region of $\hat{\mathbf{I}}_i$ shared with $\hat{\mathbf{I}}_j$. $\bar{w}_{ij}$ is a normalized weight proportional to the area of the overlap between $\mathbf{I}_i$ and $\mathbf{I}_j$ ($\sum \bar{w}_{ij} = 1$). $\triangle \mathrm{H}(\bullet)$ calculates the difference between histograms of images by bins, while $\triangle \mathrm{Go}(\bullet)$ computes the difference between gradient orientation maps of images by pixels. N_{bin} is the number of the bins of a histogram, and N_{pix} is the amount of pixels of $\mathbf{I}_i$.

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

Figure 5: Color correction results of LUNCHROOM: edited images (a), corrected results of [18] (b), [15] (c) and ours (d). For comparison's sake, regions with tonal inconsistency and color cast are marked with red boxes and green boxes respectively.

5.1. Strip Panorama of Ground Scenes

We first demonstrate the performance of our color correction method on a public panorama dataset LUNCHROOM (selected 15 images). Since the original images have little tonal difference, we deliberately modify their color CDF independently in *Photoshop*, in order to test the algorithms' robustness under drastic color and illumination variations. The edited images and corrected results are shown in Figure 5 in the form of aligned panorama. Here, to compare color disparity between adjacent images, we stitch images into a panorama with no boundary fusion applied. Inspecting the results qualitatively, we see that despite the huge input variability, the color and luminance consistency improved greatly on the corrected images. In Figure 5, we can find (d) has the best color consistency, which is also attested in Table 1. Also, on overall visual effect, (d) is slightly superior to (c) and obviously better than (b) that shows kind of dark tone and residual color cast. Our flexible correction model and energy term considering dynamic range make effects here. As for gradient preservation, Table 1 shows [15] has the best performance. This might be explained by that its albedo reflect based correction model is more close to the real imaging principle. Due to the curve slop constraint

Table 1: Quantitative comparison on color corrected results obtained through different methods. Color discrepancy (CD), gradient loss (GL) in Eq. (9) and running time (unit: second) are evaluated comprehensively.

Methods	LUNCHROOM			CAMPUS			ZY-3			UAV		
	#CD	#GL	#Time	#CD	#GL	#Time	#CD	#GL	#Time	#CD	#GL	#Time
Input	24.41	0.00	0.00	13.69	0.00	0.00	13.00	0.00	0.00	18.08	0.00	0.00
Shen *et al.*'s [18]	6.93	1.03	3.53	5.73	0.52	3.02	4.81	0.51	8.78	3.79	0.63	231.07
Park *et al.*'s [15]	8.67	0.94	102.42	7.37	0.39	90.46	6.56	0.47	3295.1	5.21	1.13	9531.4
Our approach	4.24	1.00	2.89	4.08	0.43	2.15	2.80	0.50	7.76	1.38	0.64	217.80

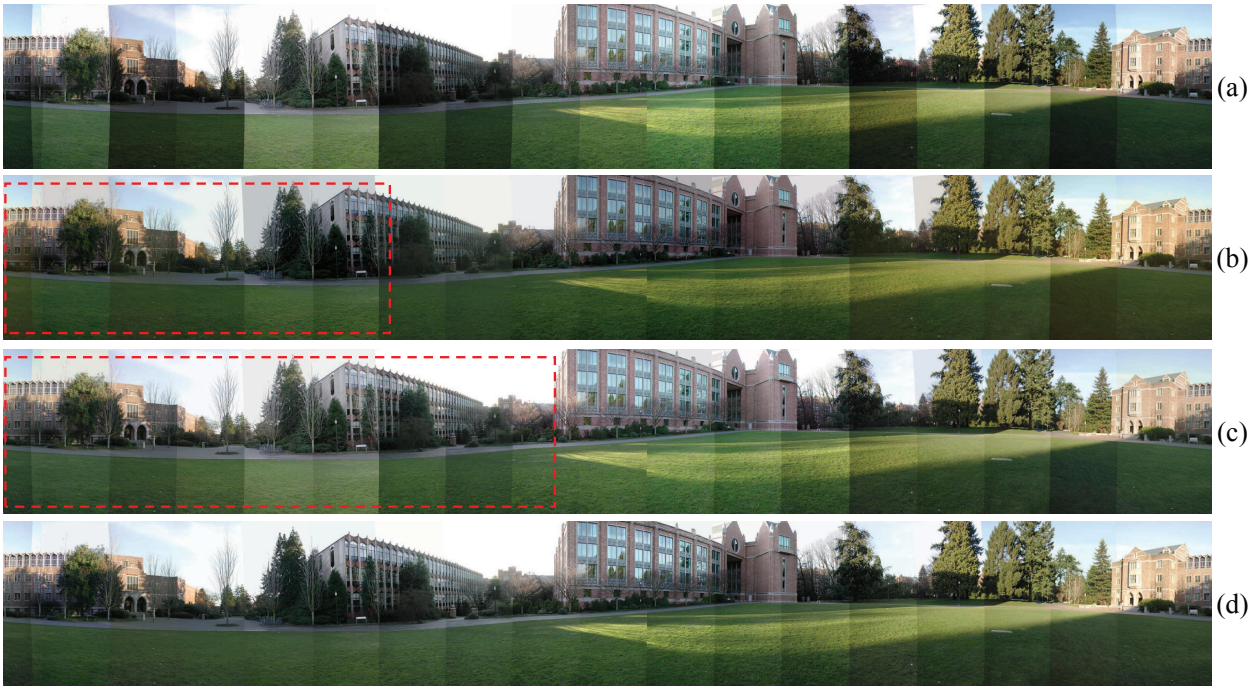


Figure 6: Color correction results of CAMPUS: source images (a), corrected results of [18] (b), [15] (c) and ours (d). For comparison's sake, regions with noticeably residual tonal difference are marked with red boxes.

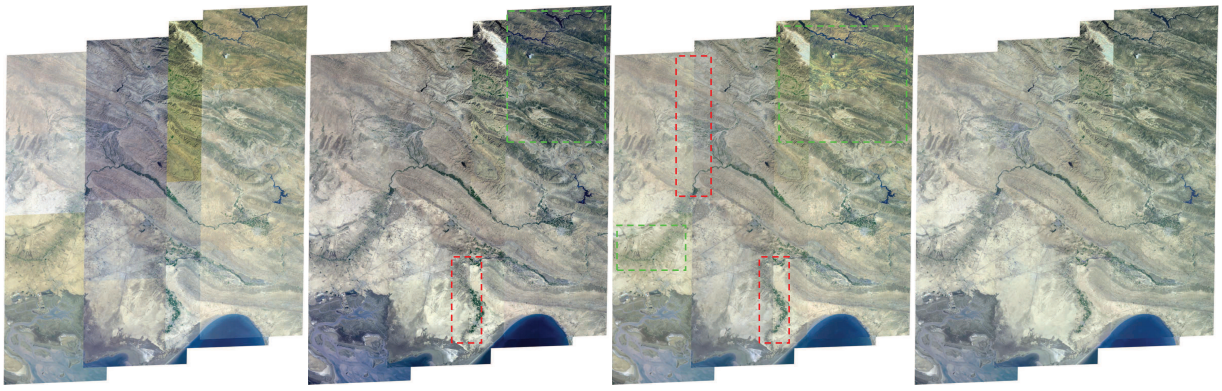
preventing intensity merging, our method outperforms linear model based [18] in gradient preservation slightly.

We further test our method on another public dataset CAMPUS (18 images) with continuous luminance variation. The source images and their corrected results are displayed in Figure 6. It is easy to observe that (d) shows the best tonal consistency, illustrating the outstanding ability of our method in minimizing color disparity. Since little color cast exists in the original images, the corrected result (b) of [18] shows better consistency than (c) of [15], such as the regions marked with red boxes in Figure 6. Note that although no approach removes the tonal difference from images completely, the residual inconsistence after correction will be low enough to conceal through fusion algorithms.

5.2. Block Mosaic of Remote Sensing Images

Stitching block-layout images usually happens in remote sensing image mosaicking, where each source image has more neighbors than in strip panorama. In some occasion, we need to correct the color disparity of images acquired at different times (even different seasons) to show consistent tone. Here, we experiment on 16 multi-temporary images acquired by Chinese ZY-3 satellite, which present obvious color disparity as shown in Figure 7 (a). The corresponding corrected results are depicted as (b), (c) and (d) in Figure 7, from which we can find that (d) has an obvious superiority over (b) and (c) in tonal consistency. In contrast, (b) has some overly dark regions and (c) has residual yellow tone uncorrected (marked in green boxes) and noticeable tonal difference between image strips. In [18], the intensity of all the corrected images will be stretched linearly to the gamut by fixed coefficients after global optimization, which turns (b) into the highest contrast but meanwhile darkens some regions overly. This might also be the reason that [18] has less gradient loss than our method in dataset ZY-3.

Further on, we analyze the performance of our method on a larger dataset UAV (130 images) captured by unmanned aerial vehicle. Similar to LUNCHROOM, the color differ-



(a) (b) (c) (d)

Figure 7: Color correction results of ZY-3: source images (a), corrected results of [18] (b), [15] (c) and ours (d). For comparison's sake, regions with tonal inconsistency and color cast are marked with red boxes and green boxes respectively.

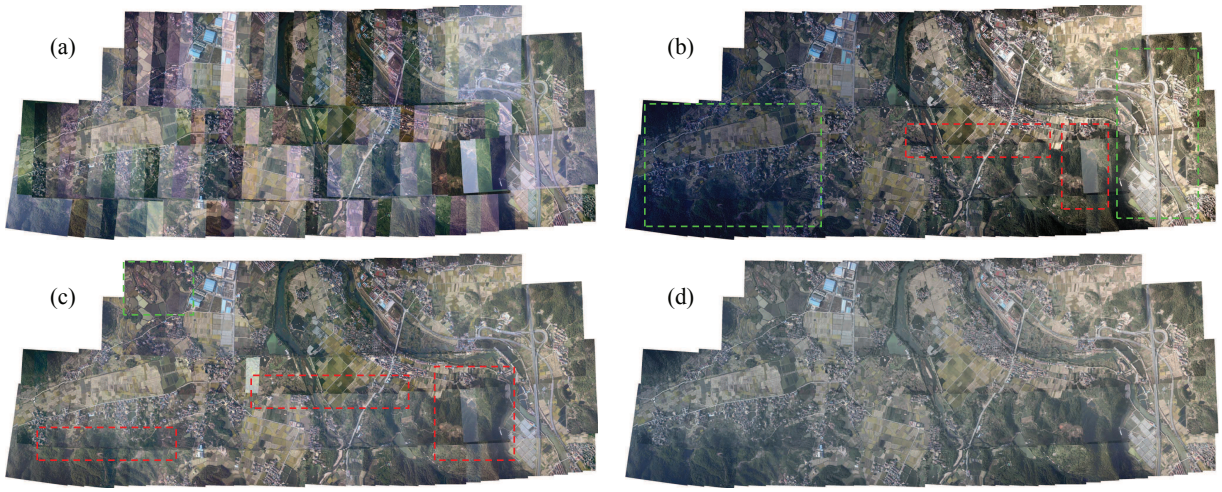


Figure 8: Color correction results of UAV: edited images (a), corrected results of [18] (b), [15] (c) and ours (d). For comparison's sake, regions with color difference and abnormal tone are marked with red boxes and green boxes respectively.

ences among images in UAV are amplified via our random tone procedure. The correction results are compared in Figure 8. Because of the global optimization scheme, all the methods exhibit little performance declining on the large dataset, except for (b) in Figure 8 showing dramatic luminance polarization. This is because [18] only considers the inter-image color disparity but neglect constraint on individual image's dynamic range. Differently, our optimization function has a special energy term that penalizes each image's dynamic range to narrow down, which makes good effect in addressing this problems. As (c) shows, [15] has a good effect in color saturation, despite of its inferior tonal consistency. Besides, some source images of low-quality gradient structures in UAV make the numerical result in gradient loss (GL) kinda abnormal and meaningless.

Conclusively, [18] shows a better ability in improve the color consistency than [15], while [15] is superior to [18] in guaranteeing good quality of corrected images, such as gradient preservation, natural contrast and color fidelity. Due to the quality-aware energy function, our method makes the best effects in generating a visually pleasing correction result.

Running Times. Our method and Shen *et al.*'s method [18] are implemented in C++, while the code of Park *et al.*'s method [15] is provided in Matlab implementation. All the procedures are tested on a desktop PC with Intel i7-2600 CPU @ 3.40GHz and 8 GB RAM, whose running time on the four datasets is depicted in Table 1. In our method, color correspondences extraction takes the major computation, and closed-form solution makes our method surpass [18] whose parameters of linear model are optimized by non-linear function. As [15] was implemented in Matlab, we turn to analyze its algorithm complexity which reflects the running efficiency in significant degree. Structure from motion (SFM) based color correspondence and its inner iterative scheme of matrix decomposition definitely make the computation cost of [15] much higher than other methods.

6. Conclusions

We have presented an effective method to optimize the color consistency across multiple images. Our key contribution lies in the novel energy function design based on the parameterized quadratic spline model, which turns major semantically visual requirements into effective parametric expression. Convex quadratic programming gives the global optimal solution quickly. Several typical experimental results verified the validity and generality of the proposed method. Comparing to existing methods, our method leads in the performance of both effectiveness and efficiency. However, single-channel optimization strategy can not solve the color cast (white balance) problem essentially. Multi-channel joint adjustment or response function calibration might be helpful solutions, which will be investigated in the future works.

Acknowledgments

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

[1] M. Brown and D. G. Lowe. Automatic panoramic image stitching using invariant features. *International Journal on Computer Vision (IJCV)*, 74(1):59–73, 2007.

[2] V. Bychkovshy, S. Paris, E. Chan, and F. Durand. Learning photographic global tonal adjustment with a database of input/output image pairs. In *CVPR*, 2011.

[3] O. Frigo, N. Sabater, V. Demoulin, and P. Hellier. Optimal transportation for example-guided color transfer. In *Asian Conference on Computer Vision (ACCV)*, 2014.

[4] B. Gooch, M. Ashikhmin, E. Reinhard, and P. Shirley. Color transfer between images. *IEEE Computer Graphics and Applications (CGA)*, 21(5):34–41, 2001.

[5] Y. HaCohen, E. Shechtman, D. B. Goldman, and D. Lischinski. Non-rigid dense correspondence with applications for image enhancement. *ACM Transactions on Graphics (TOG)*, 30(4):70, 2011.

[6] Y. HaCohen, E. Shechtman, D. B. Goldman, and D. Lischinski. Optimizing color consistency in photo collections. *ACM Transactions on Graphics (TOG)*, 32(4):38, 2013.

[7] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *CVPR*, 2016.

[8] Y. Hwang, J.-Y. Lee, I. S. Kweon, and S. J. Kim. Color transfer using probabilistic moving least squares. In *CVPR*, pages 3342–3349, 2014.

[9] V. Laparra, A. Berardino, J. Ballé, and E. P. Simoncelli. Perceptually optimized image rendering. *arXiv preprint*, page arXiv:1701.06641, 2017.

[10] L. Li, J. Yao, X. Lu, J. Tu, and J. Shan. Optimal seamline detection for multiple image mosaicking via graph cuts. *ISPRS Journal of Photogrammetry and Remote Sensing*, 113:1–16, 2016.

[11] L. Li, J. Yao, R. Xie, M. Xia, and W. Zhang. A unified framework for street-view panorama stitching. *Sensors*, 17(1):1, 2016.

[12] R. M. H. Nguyen, S. J. Kim, and M. S. Brown. Illuminant aware gamut-based color transfer. *Computer Graphics Forum (CGF)*, 33(7):319–328, 2014.

[13] A. A. Nielsen. The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data. *IEEE Transactions on Image Processing (TIP)*, 16(2):463–478, 2007.

[14] J. Pan, M. Wang, D. Li, and J. Li. A network-based radiometric equalization approach for digital aerial orthoimages. *IEEE Geoscience and Remote Sensing Letters (GRSL)*, 7(2):401–405, 2010.

[15] J. Park, Y.-W. Tai, S. N. Sinha, and I. S. Kweon. Efficient and robust color consistency for community photo collections. In *CVPR*, 2016.

[16] F. Pitié, A. C. Kokaram, and R. Dahyot. Automated colour grading using colour distribution transfer. *Computer Vision and Image Understanding (CVIU)*, 107(1-2):123–137, 2007.

[17] Y. Qian, D. Liao, and J. Zhou. Manifold alignment based color transfer for multiview image stitching. In *IEEE International Conference on Image Processing (ICIP)*, pages 1341–1345, 2013.

[18] T. Shen, J. Wang, T. Fang, and L. Quan. Color correction for image-based modeling in the large. In *Asian Conference on Computer Vision (ACCV)*, 2016.

[19] N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: Exploring photo collections in 3D. *ACM Transactions on Graphics (TOG)*, 25(3):835–846, 2006.

[20] Z. Su, D. Deng, X. Yang, and X. Luo. Color transfer based on multiscale gradient-aware decomposition and color distribution mapping. In *ACM International Conference on Multimedia (MM)*, pages 753–756, 2012.

[21] Z. Su, K. Zeng, L. Liu, B. Li, and X. Luo. Corruptive artifacts suppression for example-based color transfer. *IEEE Transactions on Multimedia (TMM)*, 16(4):988–999, 2014.

[22] R. Szeliski, M. Uyttendaele, and D. Steedly. Fast poisson blending using multi-splines. In *IEEE International Conference on Computational Photography (ICCP)*, 2011.

[23] Y.-W. Tai, J. Jia, and C.-K. Tang. Local color transfer via probabilistic segmentation by Expectation-Maximization. In *CVPR*, pages 747–754, 2005.

[24] B. Wang, Y. Yu, and Y.-Q. Xu. Example-based image color and tone style enhancement. *ACM Transactions on Graphics (TOG)*, 30(4):64, 2011.

[25] F. Wu, W. Dong, Y. Kong, and X. Zhang. Content-based colour transfer. *Computer Graphics Forum (CGF)*, 32(1):190–203, 2013.

[26] M. Xia, J. Yao, L. Li, R. Xie, and Y. Liu. Consistent tonal correction for multi-view remote sensing image mosaicking. In *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2016.

[27] M. Xia, J. Yao, R. Xie, L. Li, and W. Zhang. Globally consistent alignment for planar mosaicking via topology analysis. *Pattern Recognition*, 66:239–252, 2017.

[28] X. Xiao and L. Ma. Gradient-preserving color transfer. *Computer Graphics Forum (CGF)*, 28(7):1879–1886, 2009.

[29] Y. Xiong and K. Pulli. Color matching of image sequences with combined gamma and linear corrections. In *International Conference on ACM Multimedia (MM)*, pages 261–270, 2010.

[30] W. Xu and J. Mulligan. Performance evaluation of color correction approaches for automatic multi-view image and video stitching. In *CVPR*, pages 263–270, 2010.

[31] J. Yao and W.-K. Cham. An efficient image-based rendering method. In *Proceedings of the International Conference on Pattern Recognition*, 2004.

[32] J. Yao and W.-K. Cham. Image-based modeling and rendering from multiple views: A multiple cues based propagation approach. In *IEEE International Conference on Image Processing*, 2005.