GLOBALLY CONSISTENT ALIGNMENT FOR MOSAICKING AERIAL IMAGES

Menghan Xia, Jian Yao†, Li Li and Xiaohu Lu

School of Remote Sensing and Information Engineering, Wuhan University, P.R. China
†Email: jian.yao@whu.edu.cn

ABSTRACT

In this paper, we present a robust method to efficiently create a globally consistent and seamless mosaic from aerial images. Firstly, a globally consistent registration strategy is proposed to align the aerial images in a common coordinate system, which combines the affine model with the homographic model effectively. To suppress the accumulation of perspective distortions induced by a sequential set of aerial images taken from a wide-range region, we proposed to initially align each image by an affine model and then perform a homographic refinement in groups to increase the global consistency. Secondly, to efficiently conceal the parallax between aligned images in overlap regions with large depth differences where it is impossible to recover a highly accurate consistent image registration, a novel optimized seamline detection algorithm in the graph cuts energy minimization framework is proposed to find optimal seamlines within overlap regions for image mosaicking through rounding visually obvious foreground objects. Finally, experimental results on several representive image sets illustrate the superiority of our proposed approaches.

Index Terms— Image Mosaicking, Seamline Detection, Image Parallax, Global Consistency

1. INTRODUCTION

Nowadays, satellite and aerial remote sensing are common technique ways to quickly capture images for territorial monitoring in both civil and military fields, such as natural disaster monitoring and ground reconnaissance. Due to the range and resolution limitation of imaging sensors, a single image frame is too small to cover a wide-range ground area. Therefore, the robust image mosaicking is a reasonable technique to be applied to stitch sequential images to create a wide-view seamless mosaic image. Although there are many related studies in the literature in the past decade, the performance considering both accurate alignment and global consistency still remains to be improved.

For image seamless mosaicking, accurately aligning images into a common coordinate system is the critical first step which directly influences the mosaicking quality [1, 2, 3]. Most of image alignment approaches can be divided into two categories: area-based approaches [4, 5] and feature-based ones [6]. For aerial image mosaicking, feature-based approaches are usually applied to recover the homographic model between images [7, 8, 9] due to the fact that the ground scene can be regarded as an approximate plane observed from the aerial photographic camera. To suppress the accumulation error from sequential images, many optimization algorithms are proposed for a global alignment. A typical global optimization method is "Bundle Adjustment" [10, 11], which aims at finding an optimal solution minimizing the total reprojection error [12]. To provide a good initial solution for global optimization, Xing et al. [13] proposed to first apply the Extended Kalman Filter [14] onto the local area, and then refine all the parameters globally. Some methods [15, 16] utilized the topological structure information of images to achieve a global registration. To prevent image suffering down-scaling effect, Elibol et al. [17] proposed to optimize point positions in the mosaic frame and alignment model in an alternate iteration scheme. Although these methods can obtain the transformation models with minimum reprojection error, the globally consistent mosaic is inaccessible for them when the model is not a strict one, such as the homography for an approximately planar scene in aerial photography which usually causes a severe accumulation of perspective distortions induced by a sequential set of aerial images taken from a wide-range region , as illustrated in Fig. 1. In this case, the available global constraint can make a big difference, such as the approximate coplanarity of image planes which is utilized in our approach.

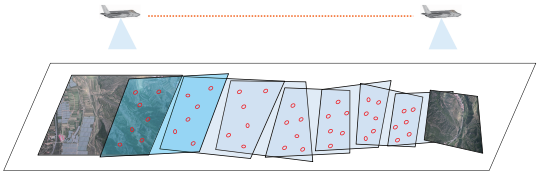

Fig. 1. An illustration of error accumulation due to perspective distortion from sequential images.

After the global alignment is recovered, the following work for creating a seamless mosaic from multiple aerial images captured from a wide-range region is to try to conceal the parallax between images induced by an inaccurate alignment inevitably existing in a region with large depth differences. Generally, the methods of concealing the parallax between images for image mosaicking [18, 19] can be divided into two main categories, smoothing transition and optimal seamline detection. Smoothing transition methods try to make given seams invisible and remove stitching artifacts by smoothing color differences between input source images. Alpha blending [20] is a simple and fast smoothing transition method, and has been widely used in image mosaicking. Optimal seamline detection methods search for the seamlines in overlap regions between images, around which intensity or gradient differences are not obvious in vision. Many methods regarded the seamline detection as an energy optimization problem and solved it by minimizing the special energy function which is defined to represent the image difference along the seamlines [21, 22].

In this paper, we proposed a robust approach to create a globally consistent and seamless mosaic from aerial images, whose flowchart is illustrated in Fig. 2. Firstly we align the aerial images into a common coordinate system via an efficient globally consistent registration strategy recovering the affine model followed by the refinement of homographic model in groups, which can magnificently suppress the error accumulation of perspective distortions induced by a

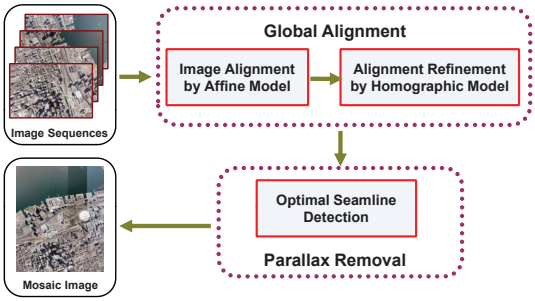

Fig. 2. The flowchart of our proposed approach for mosaicking aerial images based on global alignment and parallax removal.

sequential set of aerial images captured from a wide-range region. Secondly a novel optimized seamline detection algorithm formulated in a graph cuts energy minimization framework is performed for efficiently concealing the parallax between aligned images in regions with large depth differences where it is difficult for recovering a highly accurate consistent image registration based on the planar assumption. Finally some convictive experimental results illustrate the good performance of our proposed approach.

2. GLOBALLY CONSISTENT ALIGNMENT

To recover a robust global alignment of a large set of aerial images, especially from a single strip, it is very necessary to utilize the available geometric constraint based on the imaging condition during the global registration optimization procedure. Since all the aerial images are captured at a similar altitude and nearly parallel with the ground, all image frame planes are on a common plane approximately, from which we can infer that the transformation model between images is close to a 2D rigid transformation model. This globally geometric constraint will be very helpful to keep the global consistency via suppressing the accumulation of perspective distortions, especially for the case of the long flight strips and the case of low flight altitudes.

2.1. Image Alignment by Affine Model

To balance global consistency and alignment precision, we choose the affine model to initially align each image, as a compromise between the 2D rigid transformation and the homographic transformation. On the one hand, approximately coplanar constraint is partly implied in the six-parameter affine model which can suppress severe perspective distortion to some extent due to accumulation of errors, on the other hand, the affine transformation can provide a qualified initial solution for the following homographic refinement.

Before aligning multiple images, we have to select the reference frame which can be an image frame or an arbitrary coordinate frame. As an image reference frame, the optimal one should satisfy some criteria that its location is close to the center of the observed area and most of its corresponding feature points to adjacent frames are on the approximately planar ground. In this paper, while aligning a new image to the reference frame, all the previously aligned images having an overlap with the new image will be jointly used in the optimization framework. Let $\mathcal{I} = \{\mathbf{I}_i\}_{i=1}^{n-1}$ be a set of aligned images and $\mathbf{A}_i$ be the 2×3 affine transformation matrix of the image $\mathbf{I}_i$ with respective to the reference frame. The affine transformation $\mathbf{A}_n$ of the newly introduced image $\mathbf{I}_n$ for alignment will be optimized by minimizing the following cost function:

$$E(\mathbf{A}_n) = \sum_{i=1}^{n-1} \sum_{k=1}^{M_{i,n}} \|\mathbf{A}_i \mathbf{x}_{i,n}^k - \mathbf{A}_n \mathbf{x}_{n,i}^k\|^2, \qquad (1)$$

where $M_{i,n}$ denotes the total number of matches $\{(\mathbf{x}_{i,n}^k, \mathbf{x}_{n,i}^k)\}_{k=1}^{M_{i,n}}$ between $\mathbf{I}_i$ and $\mathbf{I}_n$, and $(\mathbf{x}_{i,n}^k, \mathbf{x}_{n,i}^k)$ denotes the homogeneous coordinates of two corresponding points from $\mathbf{I}_i$ and $\mathbf{I}_n$, respectively.

In fact, we often normalize the coordinates of matched points according to the method proposed in [23], in order to increase the numerical stability by improving the condition number of the coefficient matrix. What's more, the robust estimator MLESAC [24] is employed to exclude outliers for affine estimation, because it is beneficial for the mosaic of quasi-planar scenes. The optimal solution of $\mathbf{A}_n$ in Eq. (1) is obtained by the Singular Value Decomposition (SVD) method.

2.2. Alignment Refinement by Homographic Model

The above estimation process just recovers the locally optimal affine model of each image. To get the globally optimal estimation, the parameters of all the transformation models must be optimized jointly. Different from the method used in [25], we use the homographic models to refine the alignments of aligned images under the global consistency constraint already implied in the affine models. To increase the stability of optimization, we carry out the optimization in a two-stage strategy.

Firstly, we perform the optimization for a set of locally neighboring images as a group. Taking the single strip case for an example, the first sequential m images as a group, e.g., $\mathcal{G}_1 = \{\mathbf{I}_i\}_{i=1}^{m}$, the next group of images will be $\mathcal{G}_2 = \{\mathbf{I}_i\}_{i=m-2}^{2m-3}$, and so on. There are two common images between two adjacent groups, which mainly considers the existence of overlap in three sequential images. The homographic models of a group of images with respective to the reference frame are jointly optimized just in groups. If the homographic models of images overlapped with the previous groups have been optimized, they will be fixed in the current group. It should be noticed that a bigger value of m contributes to the global consistency while decreases the stability and efficiency of optimization. Given a group of images $\mathcal{G} = \{\mathbf{I}_i\}_{i=s}^{m+s-1}$, their homographic models $\mathcal{H} = \{\mathbf{H}_i\}_{i=s}^{m+s-1}$ with respective to the reference frame can be jointly optimized via minimizing the sum of squares of the point reprojection errors between images in $\mathcal{G}$ as:

$$E_r(\mathcal{H}) = \sum_{\mathbf{I}_p,\mathbf{I}_q \in \mathcal{G}} \sum_{k=1}^{M_{p,q}} \|\varpi(\mathbf{H}_p \mathbf{x}_{p,q}^k) - \varpi(\mathbf{H}_q \mathbf{x}_{q,p}^k)\|^2, \quad (2)$$

where $M_{p,q}$ denotes the total number of matches $\{(\mathbf{x}_{p,q}^k, \mathbf{x}_{q,p}^k)\}_{k=1}^{M_{p,q}}$ between $\mathbf{I}_p$ and $\mathbf{I}_q$, $(\mathbf{x}_{p,q}^k, \mathbf{x}_{q,p}^k)$ denotes the homogeneous coordinates of two corresponding points from $\mathbf{I}_p$ and $\mathbf{I}_q$, respectively, and $\varpi(\mathbf{x}) = [x/w, y/w]^\top$ where $\mathbf{x} = [x, y, w]^\top$. The parameters of estimated affine models are used as initial solution for optimizing the homographic models. If the homographic model of some image in $\mathcal{G}$ has been optimized in previous groups, it will be fixed in the current optimization.

Another optimization objective is to keep the global consistency for suppressing severe perspective distortions induced by a sequential set of aerial images. In other words, the optimal homographic transformation should be close to the initially estimated affine transformation, which can be expressed as the displacements of the warped features from their initial positions in the following cost function:

$$E_d(\mathcal{H}) = \sum_{\mathbf{I}_p,\mathbf{I}_q \in \mathcal{G}} \sum_{k=1}^{M_{p,q}} \Big(\|\varpi(\mathbf{H}_p \mathbf{x}_{p,q}^k) - \mathbf{A}_p \mathbf{x}_{p,q}^k\|^2 + \|\varpi(\mathbf{H}_q \mathbf{x}_{q,p}^k) - \mathbf{A}_q \mathbf{x}_{q,p}^k\|^2 \Big), \quad (3)$$

where $\mathbf{A}_p$ and $\mathbf{A}_q$ denote the initially estimated affine models for $\mathbf{I}_p$ and $\mathbf{I}_q$ with respective to the reference frame, respectively. So

far, the cost functions defined in Eq. (2) and Eq. (3) can be linearly combined to define the final cost function:

$$E(\mathcal{H}) = E_r(\mathcal{H}) + \lambda E_d(\mathcal{H}), \tag{4}$$

where λ denotes the weight coefficient for the term E_d, which should be set to an appropriate small value because this constraint is not so strict.

After the optimization of all groups is sequentially finished, the homographic models of all the aerial images as a whole group will be further jointly optimized based on the same cost function in the second optimization stage. The Levenberg-Marquard (LM) optimization method is employed for finding the optimal solution.

3. PARALLAX REMOVAL

To remove the parallax between images in overlap regions with large depth differences for image mosaicking, we proposed an effective optimal seamline detection algorithm via graph cuts, which tries to find seamlines with smallest difference in color and gradient magnitude between images in the overlap regions. The energy cost $C(\mathbf{p})$ of a pixel $\mathbf{p}$ of an image pair $\mathcal{I} = (\mathbf{I}_p, \mathbf{I}_q)$ with an overlap is defined as:

$$C(\mathbf{p}) = C_c(\mathbf{p}) + C_g(\mathbf{p}), \tag{5}$$

where $C_c(\mathbf{p})$ and $C_g(\mathbf{p})$ represent the color difference term and the gradient magnitude term, respectively.

The color difference for the pixel $\mathbf{p}$ in two images is computed in the HSV color space rather than in RGB as:

$$C_c(\mathbf{p}) = w_h|H_p(\mathbf{p}) - H_q(\mathbf{p})| + w_s|S_p(\mathbf{p}) - S_q(\mathbf{p})|, \tag{6}$$

where $H_p(\mathbf{p})$ and $S_p(\mathbf{p})$ denote the values of H and S channels of $\mathbf{p} \in \mathbf{I}_p$ and there are the same meanings for $H_q(\mathbf{p})$ and $S_q(\mathbf{p})$. The weight coefficients w_h and w_s are used to balance the influence of the differences at the H and S channels, which were set as $w_h = 1$ and $w_s = 0.1$ in this paper, respectively.

The gradient magnitudes of each pixel in the horizontal and vertical directions are obtained via the Sobel operator in the grayscale space. The gradient magnitude cost term $C_g(\mathbf{p})$ is defined as:

$$\begin{aligned} C_g(\mathbf{p}) \;=\; & |G_p^x(\mathbf{p}) - G_q^x(\mathbf{p})| + |G_p^y(\mathbf{p}) - G_q^y(\mathbf{p})| + \\ & \alpha\left(|G_p^x(\mathbf{p})| + |G_q^x(\mathbf{p})| + |G_p^y(\mathbf{p})| + |G_q^y(\mathbf{p})|\right), \end{aligned} \tag{7}$$

where α is the balanced coefficient which was set as 0.25 in this paper, $G_p^x(\mathbf{p})$ and $G_p^y(\mathbf{p})$ denote the horizontal and vertical gradient magnitudes of $\mathbf{p} \in \mathbf{I}_p$, and there are the same meanings for $G_q^x(\mathbf{p})$ and $G_q^y(\mathbf{p})$.

We formulate the optimal seamline detection as an energy minimization problem and use graph cuts to find the solution. The energy cost $E(\mathcal{I})$ comprises of the data energy term $E_{data}(\mathcal{I})$ and the smooth energy term $E_{smooth}(\mathcal{I})$, e.g., $E(\mathcal{I}) = E_{data}(\mathcal{I}) + E_{smooth}(\mathcal{I})$, where the data energy term represents the sum of energy costs for individual pixels within $\mathbf{I}_p$ or $\mathbf{I}_q$ as:

$$E_{data}(\mathcal{I}) = \sum\nolimits_{\mathbf{p}\in\mathbf{J}} (R_{\mathbf{p}}(\mathbf{I}_p) + R_{\mathbf{p}}(\mathbf{I}_q)), \tag{8}$$

where $\mathbf{J}$ is the last mosaic image, $R_{\mathbf{p}}(\mathbf{I}_p)$ and $R_{\mathbf{p}}(\mathbf{I}_q)$ represent the cost of assigning $\mathbf{p}$ with $\mathbf{I}_p$ and $\mathbf{I}_q$, which are defined as $R_{\mathbf{p}}(\mathbf{I}_k) = 0$ if $\mathbf{p} \in \mathbf{I}_k$ otherwise $R_{\mathbf{p}}(\mathbf{I}_k) = \infty$ if $\mathbf{p} \notin \mathbf{I}_k$, $k = p, q$. According to the above definition, for each pixel $\mathbf{p}$, its data energy only depends on whether it is inside the valid region of one image.

The smooth energy term $E_{smooth}(\mathcal{I})$ is defined as:

$$E_{smooth}(\mathcal{I}) = \sum\nolimits_{(\mathbf{p},\mathbf{q})\in\mathcal{N}(\mathbf{p})} \sigma_{p,q} \cdot E_{smooth}(\mathbf{p}, \mathbf{q}), \tag{9}$$

Fig. 3. An example of optimal seamline detection for parallax removal: overlapped edges (Left) and optimal seamlines (Right).

where $\mathcal{N}(\mathbf{p})$ denotes the 4-neighborhood pixel set of $\mathbf{p}$, and the coefficient $\sigma_{p,q} = 0$ if the labels of the pixels $\mathbf{p}$ and $\mathbf{q}$ are the same otherwise $\sigma_{p,q} = 1$. $E_{smooth}(\mathbf{p}, \mathbf{q})$ represents the smooth energy between two pixels $\mathbf{p}$ and $\mathbf{q}$ on images $\mathcal{I}$, which is defined as follow:

$$E_{smooth}(\mathbf{p}, \mathbf{q}) = C(\mathbf{p}) + C(\mathbf{q}), \tag{10}$$

where $C(\mathbf{p})$ and $C(\mathbf{q})$ are the energy costs of the pixels $\mathbf{p}$ and $\mathbf{q}$ defined in Eq. (5).

Based on the above proposed optimal seamline detection via a graph cuts, the extracted seamlines in overlap regions between images can visually conceal the parallax. Fig. 3(Right) shows the seamline detection result of our proposed algorithm on 6 images of a single strip [1], where our optimal seamlines successfully bypassed the tall buildings where there exist large parallaxes between images. The use of optimal seamlines can create a more coherent mosaic image compared with the case of no seamlines. Some enlarged parts below the mosaic image shown in Fig. 3 illustrates the validity of our proposed optimal seamline detection algorithm.

4. EXPERIMENTAL RESULTS

To sufficiently test the performance of our proposed algorithm for global alignment of aerial images, we chose a set of 90 aerial images in three strips with the down-sampling image size of 746×480 and a set of 48 UAV images in three strips with the down-sampling image size of 800×533 for evaluation. These two test image sequences are mainly different in flight altitudes and stability of image acquisition platforms, in which there exist an overlapping rate of about 60% in adjacent images.

4.1. Evaluation on Selection of Initial Models

In the period of recovering initial alignment described in Section 2.1, the selection of the transformation model among *rigid*, *affine* and *homographic* models can make some differences to the final mosaicking result, which are compared through conducting experiments on the two datasets above. The comparative analysis was made of both alignment precision and global consistency, and the numerical results are shown in Table. 1 ("GO" denotes the final global optimization) while the global consistency can be judged via the results shown in Fig. 4.

As for the aerial image sequence, the homographic model chosen as the initial model has the best alignment precision, but suffers severe error accumulation of perspective distortions, as shown in the last row of Fig. 4, because a homographic model has eight free parameters for alignment without the consideration of the globally consistent constraint. However, although a little inferior to that of the homographic model in alignment precision, the rigid or affine model

[1] Available at http://www2.isprs.org/commissions/comm3/wg4/tests.html

Table 1. RMS of the reprojection errors for different transformation models chosen for initial alignment (Unit: pixel).

Models	Aerial Images		UAV Images	
	RMS	RMS (GO)	RMS	RMS (GO)
Rigid	0.848	0.733	2.033	1.476
Affine	0.795	0.716	1.742	1.429
Homographic	0.497	0.492	3.838	1.633

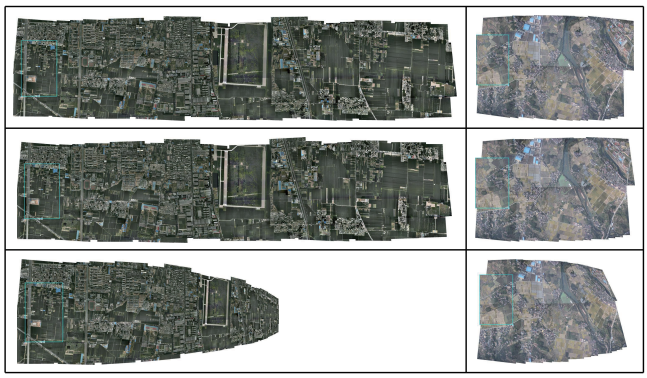

Fig. 4. The thumbnails of the mosaicking results on the aerial images (Left) and the UAV images (Right) where the rigid model in the first row, the affine model in the second row, and the homographic model in the last row were chosen for initial alignment, respectively.

chosen as the initial model has obvious superiority over the homographic model in global consistency, which benefits greatly from the approximate coplanarity implied in the characteristics of these two models. Because of the low flight altitude, the comparatively large-depth-difference ground greatly decreases the alignment precision for the UAV image sequence, especially for the homographic model. In this case, the global constraint is very useful in the optimization stabilization, which produces a globally consistent mosaic with good alignment precision. Compared with the rigid model, the affine model has a better alignment precision and a much lower computation cost in optimization, especially for the case of low flight altitude. Therefore, the affine model is the reasonable optimal choice for initial alignment. In addition, the final global optimization used in the second stage in Section 2.2 can produce higher alignment precisions as shown in Table. 1 because all the images as a whole group would be jointly optimized.

4.2. Comparative Evaluation

In this section, we compared our approach with a commercial software named PTGui [2]. The visually comparative results of mosaicking aerial images and UAV ones are illustrated in Fig. 5 and Fig. 6 respectively, in which the optimal seamline detection and image blending are not applied.

From the mosaics shown in Fig. 5, the two mosaics have a similar visual effects as a whole, though PTGui is a semi-automatic software which involves some manual adjustment for improving global consistency. However, when in alignment precision, our approach has an obvious superiority over PTGui, which can be observed from some enlarged regions. As for the UAV images, the large-depth-difference ground makes the assumption of planarity of the scene weaker, which results in even hard for keeping global consistency. The slightly down-scale tendency can be found in the mosaicking result of our approach, such as the top-left part in Fig. 6(Left). Since some strong constraints are employed for keeping the scale of each image consistent, the mosaicking result of PTGui has a quite different shape with that of our approach, meanwhile these strong constraints greatly destroyed the alignment precision, as the enlarged parts of the red box regions shown in Fig. 6. Due to the limit of pages, more experimental results and analysis are presented at `http://cvrs.whu.edu.cn/projects/MAI/`.

[2]http://www.ptgui.com/

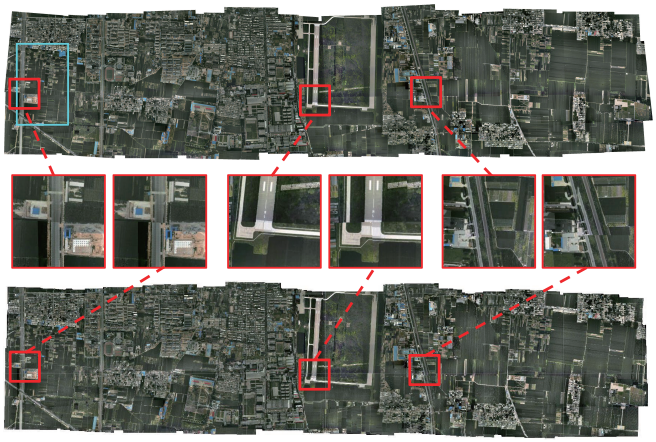

Fig. 5. The mosaics created by our approach (Top with the reference image marked in Cyan box) and PTGui (Bottom) from aerial images.

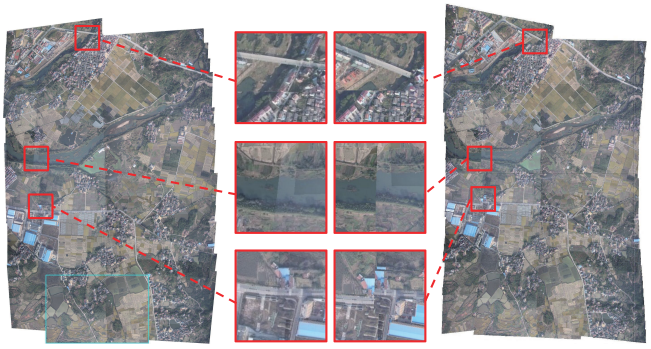

Fig. 6. The mosaics created by our approach (Left) and PTGui (Right) from UAV images.

5. CONCLUSION

In this paper, we have proposed a novel aerial image mosaicking method based on global alignment and parallax removal , which can create a globally consistent mosaic image concealing the parallax between images based on the optimal seamline detection via graph cuts. The first contribution is the effective combination of affine and homographic models in a multi-stage optimization process which keeps both global consistency and precise alignment. The second contribution is the optimal seamline detection via graph cuts based on differences in color and gradient magnitudes for concealing the parallax between aerial images. However, the global consistency is not stable enough for the case of a large-depth-difference ground relative to flight altitude, especially for the UAV platform, which will be solved in the future by studying the selection of optimal reference frame and the quick auto-organization of unordered aerial images.

6. REFERENCES

[1] Ezzeddine Zagrouba, Walid Barhoumi, and Slim Amri, "An efficient image-mosaicing method based on multifeature matching," *Machine Vision and Applications*, vol. 20, no. 3, pp. 139–162, 2009.

[2] Jinwei Chen, Huajun Feng, Kecheng Pan, Zhihai Xu, and Qi Li, "An optimization method for registration and mosaicking of remote sensing images," *Optik - International Journal for Light and Electron Optics*, vol. 125, no. 2, pp. 697–703, 2014.

[3] Barbara Zitova and Jan Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.

[4] Durga Patidar and Akshay Jain, "Automatic image mosaicing: an approach based on FFT," *International Journal of Scientific Engineering and Technology*, vol. 1, no. 1, pp. 01–04, 2011.

[5] Sherin Ghannam and A. Lynn Abbott, "Cross correlation versus mutual information for image mosaicing," *International Journal of Advanced Computer Science and Applications (I-JACSA)*, vol. 4, no. 11, pp. 94–102, 2013.

[6] A. Elibol, R. Garcia, O. Delaunoy, and N. Gracias, *Efficient Topology Estimation for Large Scale Optical Mapping*, chapter A New Global Alignment Method for Feature Based Image Mosaicing, pp. 25–39, Springer, 2013.

[7] Yisong Chen, Jiewei Sun, and Guoping Wang, "Minimizing geometric distance by iterative linear optimization," in *International Conference on Pattern Recognition (ICPR)*, 2010.

[8] Wei Mou, Han Wang, Gerald Seet, and Lubing Zhou, "Robust homography estimation based on non-linear least squares optimization," in *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2013.

[9] Liwei Zhang, Youfu Li, Jianwei Zhang, and Ying Hu, "Homography estimation in omnidirectional vision under the l-norm," in *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2010.

[10] Bill Triggs, Philip F. McLauchlan, Richard I. Hartley, and Andrew W. Fitzgibbon, *Vision Algorithms: Theory and Practice, Lecture Notes in Computer Science*, chapter Bundle adjustmentła modern synthesis, pp. 298–372, Springer, 2000.

[11] Kurt Konolige, "Sparse sparse bundle adjustment," in *British Machine Vision Conference (BMVC)*, 2010.

[12] M. Li, D. Li, and D. Fan, "A study on automatic UAV image mosaic method for paroxysmal disaster," in *International Society of Photogrammetry and Remote Sensing Congress (ISPRS)*, 2012.

[13] Cheng Xing, Jinling Wang, and Yaming Xu, "A robust method for mosaicking sequence images obtained from UAV," in *International Conference on Information Engineering and Computer Science (ICIECS)*, 2010.

[14] F. Caballero, L. Merino, J. Ferruz, and A. Ollero, "Homography based Kalman filter for mosaic building. applications to UAV position estimation," in *IEEE International Conference on Robotics and Automation*, 2007.

[15] Eun-Young Kang, Isaac Cohen, and Gerard Medioni, "A graph-based global registration for 2D mosaics," in *International Conference on Pattern Recognition (ICPR)*, 2000.

[16] Roberto Marzotto, Andrea Fusiello, and Vittorio Murino, "High resolution video mosaicing with global alignment," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2004.

[17] A. Elibol, R. Garcia, O. Delaunoy, and N. Gracias, "A new global alignment method for feature based image mosaicing," *Advances in Visual Computing, Lecture Notes in Computer Science*, vol. 5359, pp. 257–266, 2008.

[18] Che-Han Chang, Yoichi Sato, and Yung-Yu Chuang, "Shape-preserving half-projective warps for image stitching," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.

[19] Fan Zhang and Feng Liu, "Parallax-tolerant image stitching," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.

[20] Matthew Uyttendaele, Ashley Eden, and Richard Skeliski, "Eliminating ghosting and exposure artifacts in image mosaics," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001.

[21] Nuno Gracias, Mohammad Mahoor, Shahriar Negahdaripour, and Arthur Gleason, "Fast image blending using watersheds and graph cuts," *Image and Vision Computing*, vol. 27, no. 5, pp. 597–607, 2009.

[22] Martin Kerschner, "Seamline detection in colour orthoimage mosaicking by use of twin snakes," *ISPRS journal of photogrammetry and remote sensing*, vol. 56, pp. 53–64, 2001.

[23] R. I. Hartley, "In defense of the eight-point algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 6, pp. 580–593, 1997.

[24] P.H.S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 138–156, 2000.

[25] Taygun Kekec, Alper Yildirim, and Mustafa Unel, "A new approach to real-time mosaicing of aerial images," *Robotics and Autonomous Systems*, vol. 62, no. 12, pp. 1755–1767, 2014.