

IMAGE STITCHING METHOD BASED ON PROJECTIVE INTERPOLATION

Le Zhang,* Dong Ren,* Zhiyong Huang,* Shuanghui Lei,* and Chong Zhang*

Abstract

Parallax error and ghosting are serious problems in unmanned aerial vehicle (UAV) image stitching. These problems occur even more obviously when stitching together large-parallax images. In this paper, an image stitching method based on projective interpolation is proposed. First, we estimate the homography matrix from the UAV images, obtain its inverse matrix. Second, the two matrices are decomposed and interpolated into sub-matrices. Then these sub-matrices are composed into two new matrices. Finally, the image stitching is finished by warping images using the new matrices. The proposed method disperses the parallax into the two images to help reduce the parallax error. Moreover, the method introduces a seam cutting technique to preserve the geometric structure and further improve local alignment. We conducted three experiments that compared the proposed method with other state-of-the-art methods. The experimental results show that the proposed method can effectively reduce parallax error, eliminate ghosting, and achieve an overall good stitching result.

Key Words

Stitch, parallax, ghosting, decompose, interpolate

1. Introduction

Image stitching is a well-studied topic [1] that combines a set of images into a larger image with a wider field of view of the scene [2]. The primary task of image stitching is to align the input images using a global parameter transformation. In spite of their robustness, global transformations are usually not flexible enough for all types of scenes and motion. For example, a projective transformation only provides accurate alignment for planar scenes or parallax-free camera motion [2].

Most image stitching algorithms share a similar pipeline: first, the transformation or warping functions that bring the overlapping images into alignment are estimated, and then, the aligned images are composited onto a common canvas [3]. Of course, perfect alignment is

rarely achieved, and the results tend to have parallax error and ghosting. To improve results, many research efforts focus on devising better alignment [4] or compositing techniques [5].

To minimize the parallax error caused by projective transformation, Carroll *et al.* proposed the content-preserving projection [6] method. This method develops a projective transformation, adapting the contents of the images to minimize the wide-angle image alignment deviation. However, when input images do not meet the conditions of projective transformation, parallax error will occur. Gao *et al.* [7] proposed dual-homography warping to specifically deal with scenes that contain two dominant planes. The warping function is defined by a linear combination of two homography matrices with spatially varying weights. Because their wrappings are based on projective transformations without improvement, the results suffer from parallax error caused by projection. Lin *et al.* [8] proposed a smoothly varying affine stitching method to deal with images. However, their main idea is still a projective transformation, and the results have parallax error. Zaragoza *et al.* [3] proposed the as-projective-as-possible (APAP) method. It consists of a global projection, warping by local projective transformation, then the use of a moving direct linear transformation (DLT) to smoothly extrapolate local projection to non-overlapping parts. Hence, it is a global projective transformation that allows local deviation because of an unsuitable model.

To obtain better stitching results, local deviations are allowed in alignments. Fan *et al.* proposed the parallax-tolerant image stitching (PT) method [9]. It completes pre-alignment by estimating the homography matrix between images, then evaluates the quality of a homography matrix by calculating the cost of finding seams [10], [11]. These processes are repeated until the optimal homography matrix is obtained. After using the homography matrix to align the images, a seam cutting technique was introduced [12] to refine the stitching result. The method can handle ordinary parallax images well, but misalignments and artefacts occur when handling large-parallax images.

To avoid the above larger parallax error, we would like to warp the images more smoothly to reduce it. Recent studies show that appropriate decomposition and interpolation [13] of the homography matrix can warp images with a smooth transition [14]. This paper proposes a novel

* College of Computer and Information Technology, China Three Gorges University, China; e-mail: ihzlp@qq.com, rendong5227@163.com, {9558038, 184070518, 707212993}@qq.com

Corresponding author: Dong Ren

Recommended by Prof. Chaomin Luo

(DOI: 10.2316/Journal.206.2016.5.206-4769)

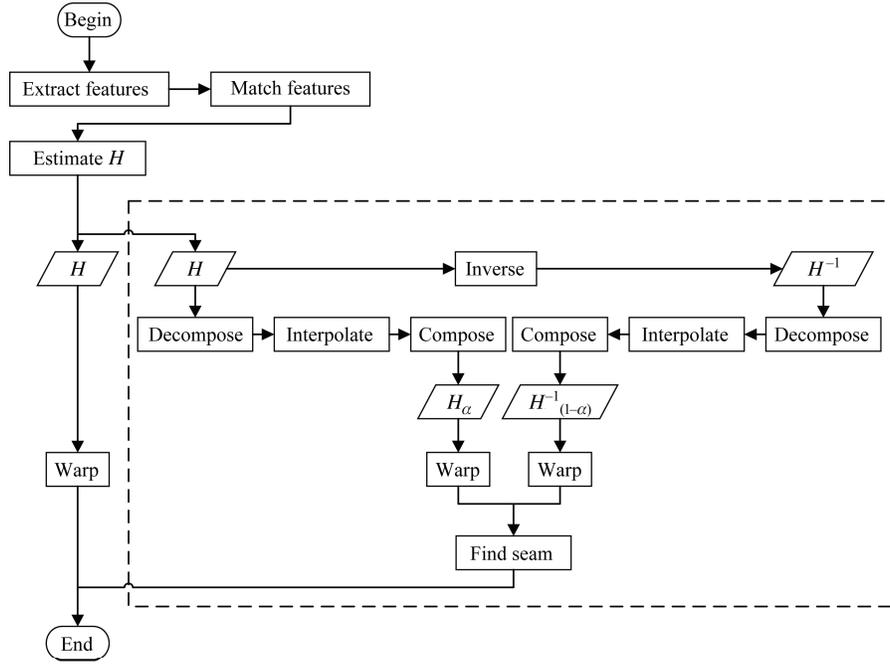


Figure 1. Block diagram of the based method and our method for stitching images.

method of image stitching based on projective interpolation. It is a combination of projective transformation and interpolation techniques. The proposed method estimates the homography and obtains its inverse matrix, then decomposes and interpolates them, respectively. Two new matrices are obtained by composing the interpolated compositions. The new matrices are then used to warp the images. Finally, a more accurate and smooth alignment is obtained. In addition, we introduce the seam cutting technique to not only ensure the alignment accuracy of the results but also to preserve the original geometry. This combination effectively disperses the initial misalignment, reduces parallax error, and obtains a good result.

2. Image Stitching Method Based on Projective Interpolation

In image stitching, the based method is using homography directly matrix to stitch images. In this way, stitching large parallax images together leads to parallax error in result. We propose a new method and compare it with the based method. The core processes of our method are in dotted box of Fig. 1. We choose the homography matrix as a projective matrix, analyse the defects of directly interpolating the homography matrix, and introduce our method for aligning images. To improve the stitching result, we introduce the seam cutting technique. An optimal seam is found to compose the final stitching results based on the results of the alignment.

2.1 Review of Projective Transformation

At present, the projective transformation of image stitching is mainly based on the homography matrix, expressed as follows:

$$p = Hq \quad (1)$$

where $p = (x, y)$ and $q = (x', y')$ are the matching points across overlapping images I_0 and I_1 , and there is a projective transformation (homography matrix) relationship between them [15].

The homography matrix is a combination of linear transformation and translation in homogeneous coordinates. Let p and q be in homogeneous coordinates, and let H be a 3×3 matrix. In homogeneous coordinates,

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \quad (2)$$

2.2 Defects of a Directly Projected Homography Matrix

Assuming that the interpolation coefficient is α ($0 \leq \alpha \leq 1$), if we linearly interpolate the matching p and q , the new point ipt is

$$ipt = (1 - \alpha)p + \alpha q \quad (3)$$

We could also directly interpolate the homography matrix, giving

$$ipt = [(1 - \alpha)I + \alpha H]p \quad (4)$$

where I is the identity matrix.

These two interpolation methods cannot obtain satisfactory results [8], because the shape interpolation and image deformation are usually done by affine interpolation on a triangle. When there is a projective transformation relationship between two images, these methods cannot yield satisfactory projective results.

2.3 Decomposition of the Homography Matrix

As shown in Fig. 1, we decompose the homography matrix into sub-matrices, and then interpolate the sub-matrices, which help to obtain a reasonable and smooth transformation.

The homography matrix H has nine unknowns. Normalizing the H , namely, dividing all elements in the matrix by h_9 , leaves only eight independent unknowns. When a matrix is multiplied by any non-zero value, the result still represents the same transformation, so we pre-multiply H by a normalization constant η ,

$$\eta = \frac{\text{sign}(h_9)}{\sqrt{h_7^2 + h_8^2 + h_9^2}} \quad (5)$$

where h_7 , h_8 , and h_9 represent the vanishing line of the planar region in homogeneous coordinates. This normalization constant η is chosen so that the decomposed projective matrix has a vanishing line vector of unit magnitude, which can avoid unnatural interpolation results.

A projective transformation can be decomposed into a chain of transformations, expressed as

$$\begin{aligned} H &= H_S H_A H_P \\ &= \begin{bmatrix} SR & t/v \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} I & t \\ V^T & v \end{bmatrix} \\ &= \begin{bmatrix} A & t \\ V^T & v \end{bmatrix} \end{aligned} \quad (6)$$

where H_S is a similarity transformation, H_A is an affine transformation, and H_P is a projective transformation. Further, K is an upper-triangular matrix normalized as $\Delta K = 1$. Matrix A is a composition of rotation and non-isotropic scaling transformations, and is a non-singular matrix given by $A = SRK + tV^T$. It can be decomposed as follows:

$$A = R(\theta)R(-\phi)SR(\phi) \quad (7)$$

According to (6) and (7), normalized projective matrix ηH can be decomposed as:

$$\eta H = TR(\theta)R(-\phi)SR(\phi)P \quad (8)$$

where $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ h_7 & h_8 & h_9 \end{bmatrix}$ is a pure projective ma-

trix extracted from the last column of ηH . Further, $V_v = \eta[h_7 \ h_8 \ h_9]^T$ is a vanishing line vector. The remainder of ηH is an affine matrix H_A , T is extracted by taking last column of H_A , and A is a sub-matrix extracted by taking the top left 2×2 part of H_A . It can be decomposed using singular value decomposition as follows:

$$A = UDV^T = (UV^T)VDV^T \quad (9)$$

According to (8), we can obtain three matrices:

$$R(\theta) = \begin{bmatrix} UV^T & 0 \\ 0^T & 1 \end{bmatrix} \quad (10)$$

$$R(\phi) = \begin{bmatrix} V^T & 0 \\ 0^T & 1 \end{bmatrix} \quad (11)$$

$$S = \begin{bmatrix} D & 0 \\ 0^T & 1 \end{bmatrix} \quad (12)$$

where $R(\theta)$ is a rotation matrix rotated by θ , $R(\phi)$ is a rotation matrix to align the axis horizontally and vertically by ϕ , and S is a scaling matrix.

2.4 Interpolation of the Homography Matrix

The left part of Fig. 1 shows how to stitch images using H . Using H directly to align the large-parallax images will result in obvious parallax error. To address this parallax error, we disperse the parallax error in the warped images into the reference and warped images by interpolating the homography matrix. As shown in the dotted box in Fig. 1, we decompose H and H^{-1} , then interpolate them to obtain two new matrices H_α and $H_{(1-\alpha)}^{-1}$. We use H_α and $H_{(1-\alpha)}^{-1}$ to warp I_0 and I_1 , respectively, to relieve parallax error. We next consider how to interpolate the homography matrix.

Assuming that I_0 is the original image and I_1 is the target image, we warp I_0 to I_1 . Image I_α is the image interpolated using coefficient α ($0 \leq \alpha \leq 1$)

$$I_\alpha = H_\alpha I_0 \quad (13)$$

$$H = T_\alpha R_{\theta_\alpha} R_{(-\phi)} S_\alpha R_{(\phi)} P_\alpha \quad (14)$$

where H_α is the interpolated homography matrix, $H_0 = I$, the identity matrix and $H_1 = \eta H$.

The progressive change from identity matrix I to projective matrix is achieved by interpolating their vanishing line vectors. The vanishing line vector of projective matrix ηH is $V_v = \eta[h_7 \ h_8 \ h_9]^T$. The vanishing line vector of I is $V_I = [0 \ 0 \ 1]^T$ in infinity. As shown in (15), the axis of rotation V_r is obtained by taking the cross product of V_v and V_I . Equation (16) shows that the cosine of rotation angle λ between the two vectors is obtained by taking the dot product of the two vectors. As expressed in (17), we interpolate λ with α to obtain the intermediate vanishing line vector in P_α . The rotations $R_{\pm\varphi}$ that align the scaling axes to the x and y directions should not be interpolated. The scaling, rotation, and translation interpolation matrices are as follows.

$$V_r = \eta[h_7 \ -h_8 \ 0]^T \quad (15)$$

$$\lambda = \cos^{-1}(\eta h_9) \quad (16)$$

$$P_\alpha = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \sqrt{\frac{h_7^2 \sin(\alpha\lambda)}{h_7^2 + h_8^2}} & \sqrt{\frac{h_8 \sin(\alpha\lambda)}{h_7^2 + h_8^2}} & \cos(\alpha\lambda) \end{bmatrix} \quad (17)$$

$$S_\alpha = (1 - \alpha)I + \alpha S \quad (18)$$

$$R_{\theta\alpha} = \begin{bmatrix} \cos(\alpha\theta) & -\sin(\alpha\theta) & 0 \\ \sin(\alpha\theta) & \cos(\alpha\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (19)$$

$$T_\alpha = (1 - \alpha)I + \alpha T \quad (20)$$

where V_r is the axis of rotation, λ is the cosine of the rotation angle, P_α is the projective matrix after interpolation, S_α is the scaling matrix, $R_{\theta\alpha}$ is the rotation matrix, and T_α is the translation matrix.

2.5 Seam Cutting

Although using the projective interpolated matrix to warp the image can yield a better alignment, it does not completely eliminate parallax error between the images. We introduce seam-cutting technique to refine stitching result. The seam not only eliminates the parallax error but also preserves the geometric structure of images.

The optimal partition of a graph can be determined by solving the maximum flow/minimum cut (min-cut/max-flow) problem. This method obtains a globally optimal seam by using a global optimal energy minimization. The cut cost is estimated by the regional cost where the seam runs through. The cost of node f is

$$E(f) = E_{\text{data}}(f) + E_{\text{smooth}}(f) \quad (21)$$

where f is a node, sink, or target, E_{data} is the data cost, and E_{smooth} is the smoothness energy.

The optimal seam search approach is a Markov Random Field labelling problem, which minimizes global energy by

$$E = \sum_p E_{\text{data}} + \beta \cdot \sum_{(p,q) \in N} E_{\text{smooth}} \quad (22)$$

where E_{data} is the data cost that reflects the saliency of pixels and helps to classify the pixels into the correct category as much as possible, E_{smooth} is smoothness energy, which measures the discontinuity of the adjacent pixels and constrains adjacent pixels from being cut as much as possible. β is a coefficient, and N represents a four-connected neighbourhood.

The data cost of each pixel is defined as the gradient at its location:

$$E_{\text{data}}(p, l_p) = -\nabla I_{(l_p)}(p) \quad (23)$$

Table 1
Experimental Methods

Method	Abbreviate
Directly stitch with H	DH
As-projective-as-possible Image stitching with moving DLT [3]	APAP
Parallax-tolerant image stitching [9]	PT
The proposed method	HI

Table 2
Experimental Images

No.	Source
1st	APAP [3] image set
2nd	PT [9] image set
3rd	UAV images taken by our team

where E_{data} decides which image gradient to choose at pixel p . The smoothness cost between two pixels p and q is defined as:

$$E_{\text{smooth}}(p, l_p, q, l_q) = |l_p - l_q| \cdot (D(p) + D(q)) \quad (24)$$

where E_{smooth} represents discontinuities between each pair of neighbouring pixels. If $I_p = I_q$, the smoothness cost is zero, else the smoothness cost is defined as the difference D of the overlapped pixel. D is:

$$D(p) = \|I_1(p) - I_2(p)\|^2 + \gamma \cdot \|\nabla I_1(p) - \nabla I_2(p)\|^2 \quad (25)$$

where γ is a coefficient.

3. Experiments and Analysis

To test the performance of the proposed method, we conducted experiments to compare the proposed method with several state-of-the-art stitching methods. Tables 1 and 2 list the stitching methods and images tested in the experiments, respectively.

From the first to third set, the image content ranges from simple to complex, and parallax ranges from small to large. We prepared three experiments to effectively verify the stitching performance of the proposed method, especially for images with large parallax.

We first conducted experiments using the four methods on the first set of images to verify the performance of the proposed method on images with simple content and small parallax. The results are shown in Fig. 2.

Figure 2(a) is the result of stitching by DH. As homography is a global transformation, warping with it directly often does not yield good alignment results. Misalignment leads to parallax error, and ghosting appears in the composite image. In the stitching result, the areas of the

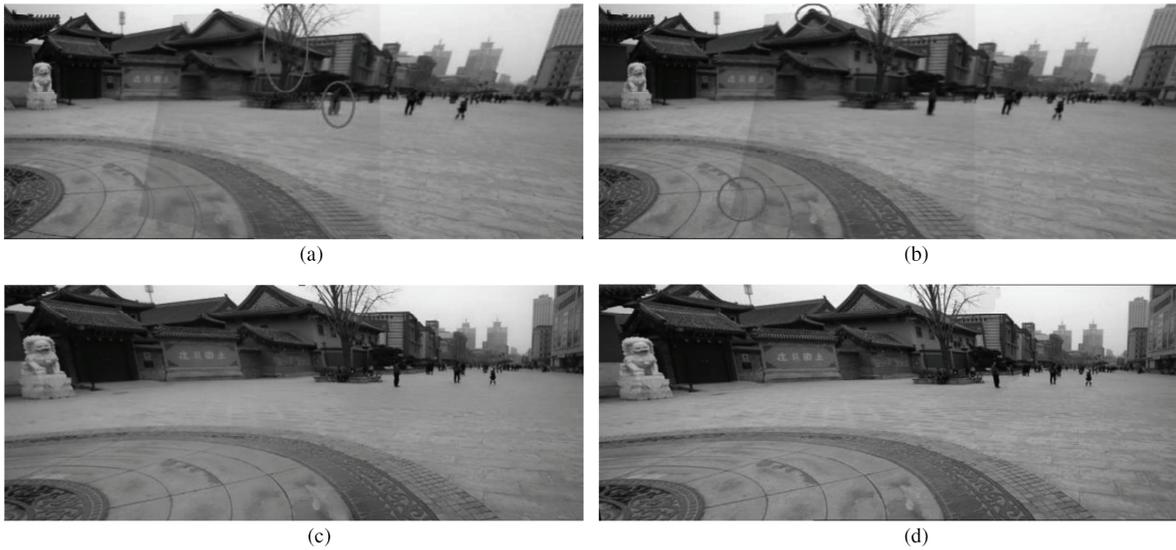


Figure 2. Comparison of various stitching method results on the APAP image: (a) DH; (b) APAP; (c) PT; and (d) HI.



Figure 3. Comparison of various stitching method results on the PT image: (a) DH; (b) APAP; (c) PT; and (d) HI.

person and trees yield ghosting. Figure 2(b) is the result of stitching by APAP, whose core idea is to improve the global homography transformation by moving DLT to adjust the homography matrix and hence relieve the image deviation. As shown in Fig. 2(b), the result has been improved, but ghosting still occurs. As Figs. 2(c) and (d) show, PT and HI effectively eliminate ghosting and yield satisfying results. The performance of the method proposed in this paper is better than DH and APAP, and similar to PT when stitching images with simple content and small parallax.

Second, we conducted experiments with the four methods on the second set of images to verify the performance of the proposed method on images with complex content and large parallax. The results are shown in Fig. 3.

Figure 3(a) shows the result of stitching by DH. Ghosting can be found in the building edges. Figure 3(b) is the result of stitching by APAP. The ghosting still

exists around the building. As a homography matrix cannot account for parallax, the result of the homography-based APAP method often yields some amount of ghosting. Figure 3(c) is the result of stitching by PT. This method is a good solution for the misalignment and ghosting caused by projective transformation, but there are parallax errors in local areas. As the result shows, the street lamp is discontinuous because of parallax error. Figure 3(d) is the result of stitching by HI. As the result shows, ghosting is eliminated and the street lamp is continuous. The HI method successfully avoids the problems of the other three methods and yields a satisfying result. The performance of the proposed method is better than that of the other three methods when stitching images with complex content and large parallax.

Third, we conducted experiments with DH, APAP, and HI on the third set images to verify the performance of proposed method on images with more complex content

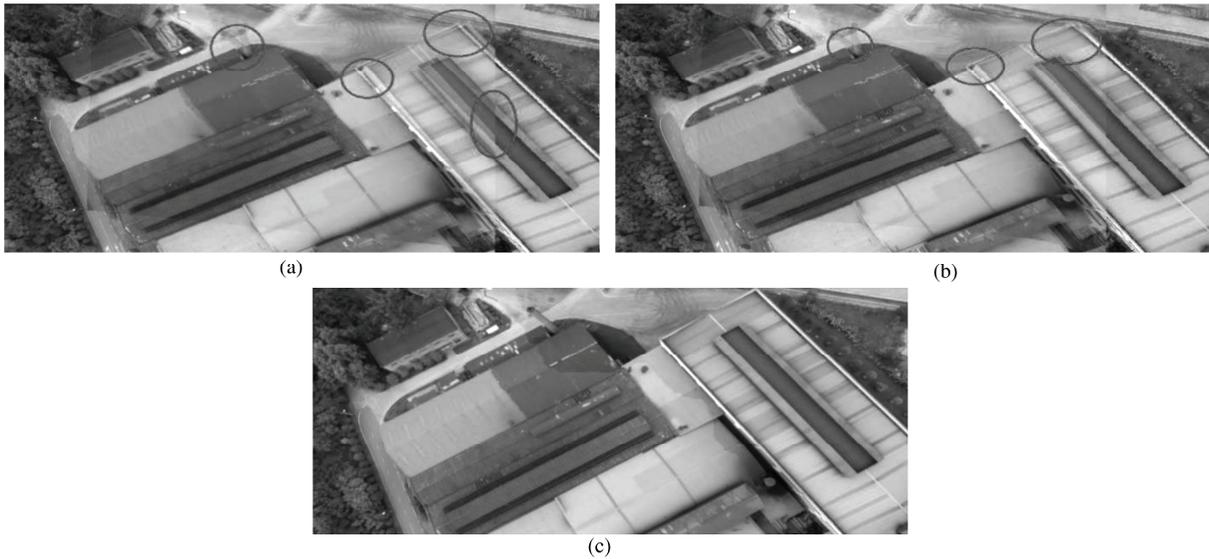


Figure 4. Comparison of various stitching method results on our image: (a) DH; (b) APAP; and (c) HI.

and an even larger parallax. The results are shown in Fig. 4.

Figure 4(a) is the result of stitching by DH. A large amount of parallax error and ghosting is found. Figure 4(b) is the result of stitching by APAP. Although parallax error and ghosting can be found in the result, it is slightly better than the results of DH. Figure 4(c) shows the result of stitching by HI. As the result shows, the ghosting around the buildings is eliminated and parallax errors are not found. The performance of the proposed method is clearly better than the DH and APAP methods when stitching images with more complex content and an even larger parallax.

From the three experiments, we can see that the results of our method do not have ghosting and parallax error, but the results of the other three methods do. The based methods usually use a homography matrix to align the images. Compared with the other three methods, our method does a large amount of work after obtaining the homography matrix. First, we decompose the homography matrix and obtain its sub-matrices. We then interpolate its sub-matrices with a coefficient α ($0 \leq \alpha \leq 1$). Finally, we compose the interpolated sub-matrices and obtain a new matrix H_α . We obtain the inverse matrix of the homography, set the interpolated coefficient as $(1 - \alpha)$ and deal with H^{-1} in the same way to obtain a new matrix $H_{(1-\alpha)}^{-1}$. We use the two new matrices to warp the images and disperse the parallax error into two images to help reduce it. Moreover, we introduce a seam cutting to preserve the geometric structure and content of the original images, further improving local alignment and obtaining a good result.

4. Conclusion

Homography-based methods can work well when the images to be stitched are obtained by rotation alone, or the image scenes are effectively planar. That stitching result

of images using the homography matrix does not meet the necessary conditions and often yields misalignment and ghosting. This paper presents a method that aligns input images with projective interpolated matrices. We estimate the homography matrix and its inverse. The two matrices are then decomposed and interpolated. Finally, we compose the interpolated elements into two new matrices, as projective interpolated matrices. As the experimental results show, the proposed method can obtain a good result on three images almost without parallax error and ghosting. The proposed method can disperse parallax to increase the accuracy and smoothness of the alignment. It is good at stitching large parallax images, especially stitching unmanned aerial vehicle (UAV) images. When stitching more than two images, we should choose a homography matrix before using proposed method to finish stitching.

Acknowledgement

This work was supported in part by the Natural Science Foundations of China (Grant no. 41371349), and National High Technology Research and Development Program of China (Grant nos. 2013AA102303 and 2014BAD04B05). This research was jointly supported by the Hi-Tech Tech Research and Development Program of China (No. 2012AA101905).

References

- [1] R. Szeliski, Image alignment and stitching: A tutorial, *Foundations & Trends[®] in Computer Graphics & Vision*, 2(1), 2006, 1–104.
- [2] C.H. Chang, Y. Sato, and Y.Y. Chuang, Shape-preserving half-projective warps for image stitching, *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, 3254–3261.
- [3] J. Zaragoza, T.J. Chin, M.S. Brown, and D. Suter, As-projective-as-possible image stitching with moving DLT, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7), 2014, 1285–1298.

- [4] J.W. Lee and I.S. Kweon, Image point feature matching by triangulation, *Intelligent Automation & Soft Computing*, 3(2), 1997, 135–149.
- [5] D. Ren, Y.M. Liu, X.D. Yang, J.H. Wang, and H.Y. Yu, An improved PCA fusion method based on generalized intensity–hue–saturation fusion technique, *Intelligent Automation and Soft Computing*, 18(18), 2012, 1165–1175.
- [6] R. Carroll, M. Agrawal, and A. Agarwala, Optimizing content-preserving projections for wide-angle images, *Acm Transactions on Graphics*, 28(3), 2009, 341–352.
- [7] J. Gao, S.J. Kim, and M.S. Brown, Constructing image panoramas using dual-homography warping, *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, 42, 2011, 49–56.
- [8] W.Y. Lin, S. Liu, Y. Matsushita, T.T. Ng, and L.F. Cheong, Smoothly varying affine stitching, *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, 32, 2011, 345–352.
- [9] F. Zhang and F. Liu, Parallax-Tolerant Image Stitching, *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Washington, DC, USA, 2014, 3262–3269.
- [10] V. Kwatra, A. Schödl, I. Essa, G. Turk, and A. Bobick, Graphcut textures: Image and video synthesis using graph cuts, *ACM Transactions on Graphics*, 22(3), 2003, 277–286.
- [11] Jr. Ford, R. Lester, and D.R. Fulkerson, Flows in networks, *Mathematics of Computation*, 18(4), 2009, 157–172.
- [12] Y. Boykov and V. Kolmogorov, An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 26(9), 2004, 1124–1137.
- [13] K. Shoemake and T. Duff, Matrix animation and polar decomposition, *Proceedings of the Conference on Graphics Interface*, 92, 1992, 258–264.
- [14] T.Y. Wong, P. Kovesei, and A. Datta, Projective transformations for image transition animations, *Image Analysis and Processing, 2007. ICIAAP 2007. 14th International Conference on IEEE*, 2007, 493–500.
- [15] H.J. Asl, G. Oriolo, and H. Bolandi, An adaptive scheme for image-based visual servoing of an underactuated UAV, *International Journal of Robotics and Automation*, 29(1), 2014, 92–104.



Dong Ren received his Ph.D. degree in Jilin University. He is a vice professor in the College of Computer and Information Technology at China Three Gorges University, Yichang, China. His research interests include artificial intelligence, pattern recognition, and 3S technology.



Zhiyong Huang received his Ph.D. degree in computer application technology from Wuhan University. He is an associate professor at the China Three Gorges University. His research interests include computer vision, computer graphics, image processing, and 3D reconstruction.



Shuanghui Lei received her B.S. degree in China Three Gorges University. She is a master's candidate of China Three Gorges University. Her main research interests are pattern recognition and image processing.



Chong Zhang is studying his M.Sc. in the College of Computer and Information Technology from China Three Gorges University. His research interests include image processing, face recognition, and 3D reconstruction based on image.

Biographies



Le Zhang received his B.S. degree in China Three Gorges University. He is now a master's candidate of Institute of Collaborative Innovation Center for Key Technology of Smart Irrigation District in Hubei, China Three Gorges University. His main research interest is UAV images stitching.