

A New Stable and Accurate Algorithm of Large Image Mosaic

Bo Cai^{1,2*}, Zhigui Liu^{2,1}, Junbo Wang^{2,1} and Yuyu Zhu²

1. China Academy of Engineering Physics, Mianyang, Sichuan 621000, China

2. Southwest University of Science & Technology, Mianyang Sichuan 621010, China

caibo_bupt@126.com

Abstract

Due to the overlapped region and image size of the input image pair is unpredictable, it makes the matching procedure more difficult and unstable. For the purpose of finding out the stable and accurate matching algorithm of large images, we give an analysis of different kinds of characters, such as, scale invariant feature transform (SIFT), local maximum gradient descriptor, Harris corners, and the maximum curvature points of the image edges, etc. Based on the experiments of different images, a new stable matching algorithm is proposed in this paper. In our model, the matching procedure is divided into two stages, the rough and accurate matching procedures. To evaluate the matching result, the edge information is combined with the local maximum gradient of the input images as the constraint of our matching algorithm. After the extraction of the local maximum gradient character points, we use the edge information to divide these points into different classes. Then, the searching of the stable and accurate matching problem becomes to find out the best matching results which agree to the edge constraints. The experimental results show that the proposed algorithm is more efficient and stable than the other kinds of matching algorithms especially in the proposing of large size images.

Keywords: Image mosaic; Harris corner; Image matching; Gradient information; Edge detector

1. Introduction

Image mosaicking is a very effective method for solving the contradiction between field of view and resolution in different vision inspection. Its principle is to utilize registration and blending techniques to stitch two or more images which have overlapped areas into one large image, which is always of wide field and highly resolution. It is well known that the detection of the characters within one image is always influenced by the perspective transform and the area of the overlapped region. When the transform matrix and overlapped regions are unknown, the detected characters may leads to wrong matching of the image pair, especially, when the image size is large, the matching procedure may become much more time-consuming.

The problem we really focus on is how to optimize the alignment error and improve the stability and efficient of the matching procedure. Because of the image contents, size, characters, and the overlapped region are unknown, the common algorithms which aimed at extracting the characters of the single image contents is the main reason of unstable matching. The wrong matching condition may be dividing into two conditions, partially matching of the detected characters of the overlapped region and the totally wrong matching of the image pair. Comparing to the totally wrong matching condition, the partially matching is the common phenomenon in the image matching processing. Only using the mutual characters of the image pair to evaluate the matching result is also an unstable algorithm because

of the self-similarity of image contents and the perspective transform of different viewpoint.

In this paper, we focus on the study of the stability of image matching procedure and the evaluation of the accurate of the matching result, especially when the image size is very large. To make the final stitching image meet human visual characteristics, in other words, the alignment error of interest area is as small as possible. The key problem is to find out the exact matching points of the overlapped region in the image pair. To make our algorithm more adaptable to the different image size, we use the two stage matching algorithm to realize the total matching procedure. In the first step, we resize the image to a small size image, and extract the character points according to the gradient or local maximum gradient. After the character points have been extracted, we roughly match the image pair, and evaluate the matching result. When the matching result meets the constraint, for example, the edge condition, the local maximum gradient orientation etc. we search the exact rough image characters according to the local regions of the matched points. And then, we calculate the accurate matching result.

The remainder of this paper is organized as follows. In section 2, we give the briefly introduction of the related work of image mosaic. The proposed matching algorithm and details of each step, such as the character point detection and edge extraction, and constraint conditions are introduced in section 3. In Section 4, we mainly give the optimization of our matching procedure. Section 5 is the experimental comparison of our model with other image matching results. We go on analyzing the results of different kinds of character detecting and matching algorithms. Finally, the conclusion and limitations of our model have been discussed in section 6.

2. Related Work

Commonly, the mosaic [1-5] of image pairs consist three steps: the extraction of the image characters, the matching of the characters, and the alignment of the image pixels. In the first step, the characters are often considered as: Harris corners, SIFT points, edge information, geometric information, high-gradient points, and some other self-constraint conditioned characters of the image pair. But, when the images are self-similarity in much of the contents, any of the above mentioned detectors may cause the wrongly matching of the detected characters. In the image matching procedure, the main purpose is to find out the character points those both existed in the left and right image. It is very difficult to find out a standard to make this step stable before matching procedure, because of the 3-dimension perspective transform is often obey the 2-dimension geometric characters. For the purpose of improving the stability of this step, some constrains, such as line, edge, geometric relation, and region character are often added to the character descriptions. On this condition, the matching of the character points becomes to the matching under the inner constrains of each image pair.





Figure 1. The Processing Image Pair

For the purpose of improving the mosaic efficient and stability, Jun-Wei Hsieh [6] use the edge pixel as the pre-matching procedure, and then use the character pixel to optimize the total matching of the input pair of images. In [7], the author use the SVM to estimate the matching and mosaic results, but, the input training images and the evaluation of the parameters is very difficult. For the purpose of improving the traditional method based on characteristic, Gang Xu et al. [8] use the granular computing algorithm to detect the image character points. Aimed at obtain the topology information, Armagan Elibol *et al.* [9] use the this information to improve the robust of their matching procedure. For the purpose of measure the different descriptors of shape, color, and size in image mosaic, Zoe Falomir [4] give a comparative analysis of different descriptors. All the above mentioned papers are aimed at improving the stability and accurate of image mosaic algorithm before the matching procedure. Aimed at improving the constraint condition, many papers have been given the study of matching based on different constraints, such as, the Delaunay triangulation and affine invariant geometric constraint [10], the template matching with arctangent Hausdorff distance measure [11], the coloring based approach for matching [12], the fast and scalable approximate spectral graph matching [13], the optimized hierarchical block matching [14], and improved robust algorithm for feature point matching [15] etc.

Given two images I_m and I_n , let $p = (x, y)^T$ be a pixel location in I_m , and $q = (x', y')^T$ be the transformed location in I_n . The affine model is mathematically expressed as:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = M * \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (1)$$

where $M = \begin{pmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & m_8 \end{pmatrix}$. The technique used to estimate M is random sample

consensus (RANSAC) [16]. RANSAC is a robust estimation procedure that uses a minimal set of randomly sampled correspondences to estimate image transformation parameters, and finds a solution that has the best consensus with the data. In the case, we

may choose the matched feature points pairs to evaluate M using the direct linear transformation (DLT) method. Unfortunately, when the input pairs of character points extracted from the first step are redundancy, this algorithm may leads to local minimum. Figure v2 shows the wrong and correct estimated results of the homograph matrix.



Figure 2. The Mosaic Result of Wrong (Left) and Correct (Right) Estimated Homograph Matrix

Under this condition, we give the deeply study of the different character points detection algorithms and the matching stability of image mosaic. Comparing to the other algorithms, the main advantages of our method can be highlighted as:

(1) For the purpose of making our algorithm more adaptable to the different contents and size of the processing images, we divided the matching procedure into two steps, the rough and accurate matching procedure.

(2) To avoid the unpredicted character point number under the threshold condition, we give a comparable study of the different character detection algorithms, and proposed a new fixed character point number detecting algorithm.

(3) Unlike the blindly pre-matching constraints extracting, we deal with the constraint condition after the rough matching procedure. It makes the evaluation of the matching result and constraint condition more easily and stable.

(4) Because of the total detecting accurate is influenced by the matching accurate of each image pairs, by adding the edge constraint and pre-matching procedure, our algorithm improved the total efficient and accurate to a large extents.

3. The Proposed Algorithm

In this section, we aimed at improving the stability and accurate of image mosaic. To improve the accurate and stability of image mosaic, the accurate of the extracted character points is very important. SIFT [17] descriptors can accurately extract scale invariant image characteristics around key points. But, when this descriptor is used in images with content self-similarity, it may lose its role. The Harris detector [18] measures the local variation by shifting in various orientations, which is convoluting the original image with a Gaussian kernel. Given the input image I , the convolution of the input image is defined as:

$$S(x, y) = G(x, y) * I(x, y) \quad (2)$$

$$G_x(x, y) = \frac{-x}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right) \quad (3)$$

$$G_y(x, y) = \frac{-y}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right)$$

where $G(x, y)$ is the Gaussian kernel function, $G_x(x, y)$ and $G_y(x, y)$ are the two directions, horizontal and vertical ones. In our paper, we use the smoothing window

of 7×7 in eq.(2) and (3) represent the positions of every pixel in the window. Therefore, two kernels G_x and G_y focusing on two orientations are formed. Then, $I_x(x, y)$ and $I_y(x, y)$ of the directional kernels convolution may be calculated.

After every convolution, we have a Th calculated as follows:

$$Th(A) = Det(A) - \epsilon Tr^2(A) \quad (4)$$

where A is matrix composed of four elements

$$A = \begin{pmatrix} a_{xx} & a_{xy} \\ a_{xy} & a_{yy} \end{pmatrix} \quad (5)$$

and a_{xx} is the summary of all the $I_x^2(x_i, y_i)$, belonging to the smoothing window. Here a_{xy} is the total value of all the $I_x(x, y)I_y(x, y)$, and accordingly, a_{yy} is the sum of $I_y^2(x, y)$.

Harris and Stephens [18] believed that the corner points are selected by examining the maximum T of every 7×7 window after generating all T in the image. And a sample image based on Harris corner detector is shown in Figure 3 (a).

Because of the Harris corner use the sized (7×7) window to calculate the maximum value of the character points, it may makes the edge information lost to a large extent. For the purpose of comparing different kinds of character points, we use the magnitude of image gradient, directional gradient of each pixels, and local maximum difference as the main characters. The magnitude of the image gradient $M_g(x, y)$ is defined as:

$$M_g(x, y) = \sqrt{G_x^2(x, y) + G_y^2(x, y)} \quad (6)$$

$$M_d(x, y) = \max \{ |G_x(x, y)|, |G_y(x, y)| \} \quad (7)$$

where $G_x(x, y)$ and $G_y(x, y)$ is the directional gradient of the image pixel in x-direction and y-direction. Figure 3 (b) shows the extracted character point based on magnitude of the image, and Figure 3 (c) is the extracted points of the directional gradient. For the purpose of extracting the local maximum difference points, we use the local minimum template of 3×3 or 5×5 filter T_L to processing the image, and then using the rough image to deduce it. The extraction of the local maximum point is based on the deduced result. Given an input image I , the local maximum difference is defined as eq.7. Figure 3 (d) is the extracted result of this kind of character points.

$$I_{LM} = I - I * T_L \quad (8)$$

where T_L is the minimum filter, when the template size is 3×3 , the calculated value of the template is $z_{i,j} = \min \{ I_{k,l}, k = i-1, i, i+1; l = j-1, j, j+1 \}$.

For the purpose of letting our algorithm more adaptable to different image size, we resized input image pair into small images, for example 128×128 . After the detection of the small size image characters, we pre-matching the two image and calculate the homograph matrix M' . Then, we use the edge information to evaluate the pre-calculated homograph matrix M' . If the homograph matrix M' was agree with the edge constraint, we choose the correct matching points as the seed pixel to search the rough image character pixel, otherwise, recalculating the homograph matrix M' . The pre-evaluated homograph matrix M' is also used as the constraining of the refining character point detection. At this step, we use the matched points to calculate the rough image local region character points. On this condition, we can avoid the calculating of the large total image. Generally, our image mosaic algorithm can be summarized in the following steps:

Step 1: Calculating and extracting the character points of the resized input image pair.

Step 2: Extracting the resized image edge and extracting the edge covered character points.

Step 3: Calculating the pre-matching homograph matrix M' .

Step 4: Checking the edge constrain condition of the rough matched character points. If the condition is matched, go to step 5, else, recalculating homograph matrix M' until the constrain condition matched.

Step 5: Calculating the accurate character points of the rough image.

Step 6: Calculating the accurate homograph matrix M .

Step 7: Transform original image and achieve built-up image in.

Step 8: Blending the input pair of images.

4. Optimization

In this section, we give the analysis of the character point extraction and the constraint condition. For more accurate mosaicking purpose, the extraction of the rough image matching points will be analyzed here too.

4.1. The Extraction of the Feature Points

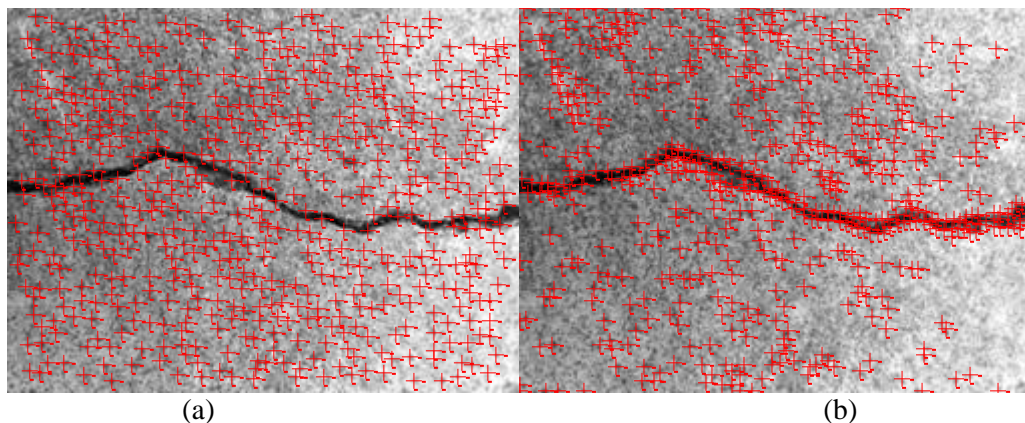
In character point extraction procedure, the number and distribution of character points are the main problems. Basically, the character points should match the following three conditions before the matching procedure.

(1) The character points should be uniformly distributed in the image, when there is any overlapped information of the image pair.

(2) The feature of the extracted points should be consistence to the global condition, instead of the local condition.

(3) To make the RANSAC processing effective and stable, the number N of the character points should be fixed or within a range.

Along with the above mentioned standards, we adjust the character point extraction algorithm. Firstly, we calculate the feature value of the resized image according to eq.(6-8). The second step is to decide the character number N . Commonly, the choosing of the character points is use a pre-decided threshold, when the threshold is higher, the detected number may be less, otherwise, the detected points may become much more. Both of the two conditions may leads to the matching procedure unstable. Considering of the strength and distribution of the image character points, we choose the maximum value of the image step by step, and after the extraction of the current maximum points, we eliminate the current points and its local (such as 5×5) adjacent points. Under this condition, we may set the number of character points previously, for example 400. The Figure .3(b-d) is the extracted character points when the number set as 400.



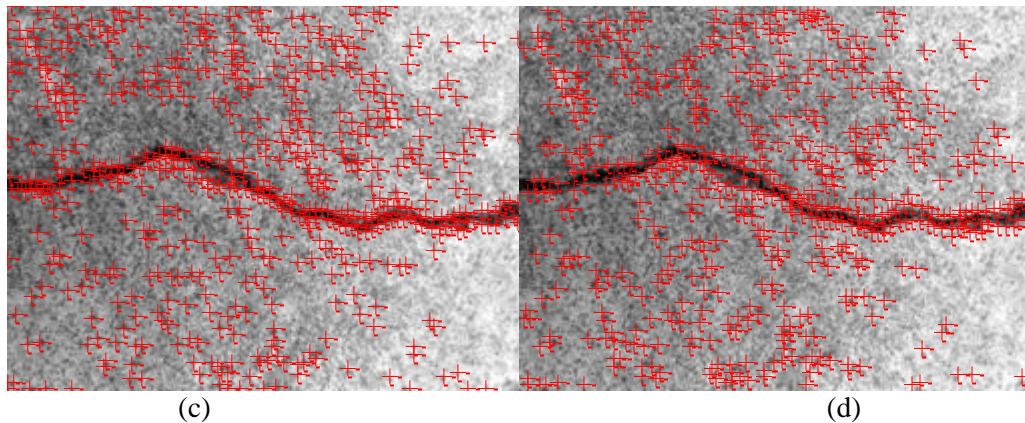


Figure 3. The Different Kinds of Extracted Character Points: (a) is the Harris Corner, (b) is the Magnitude of the Gradient, (c) is the Directional Gradient, and (d) is the Local Maximum Character Points of 5×5 Template

Comparing to the other kinds of character point extraction algorithm, the proposed algorithm may avoid the concentration of the strong response of local regions. And at the same time, it may also keep the extracted points uniformly or nearly uniformly distributed.

4.2. The classification of the Extracted Points

There are many kinds of edge detection algorithms, such as, the simple Gradient vector function, the normal Gradient Modulus function, and the Wavelet Modulus function[19], and some edge detecting templates. Finding out the enclosed region of the image is very difficult, and it is always the problem of image segmentation. Here, we need only to find out the strong line or points of the image edges. So, we choose the Sobel edge detector as the edge extraction operator. After the extraction of the main edges, we choose the extracted character points which cover the edges as the constraint condition. Fig.4 shows the result of edge covered character points.

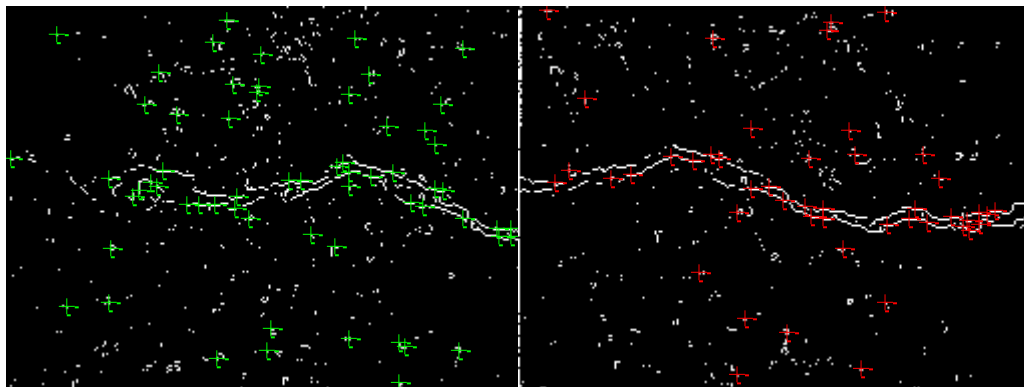


Figure 4. The Extracted Edge Covered Character Points of One Pair of Images

To measure the proposed constraint, we need to find out the accurately matching of the detected edge covered points. But, the single image based edge covered character points detector may not keep the consistence of the matching image, because of the common contents of the images are unknown before the matching procedure. The searching of the matching point in another image is the only way we can acquire them. Fortunately, the

pre-matched homograph matrix M' has provided us the information of the roughly searching region in the matched image.

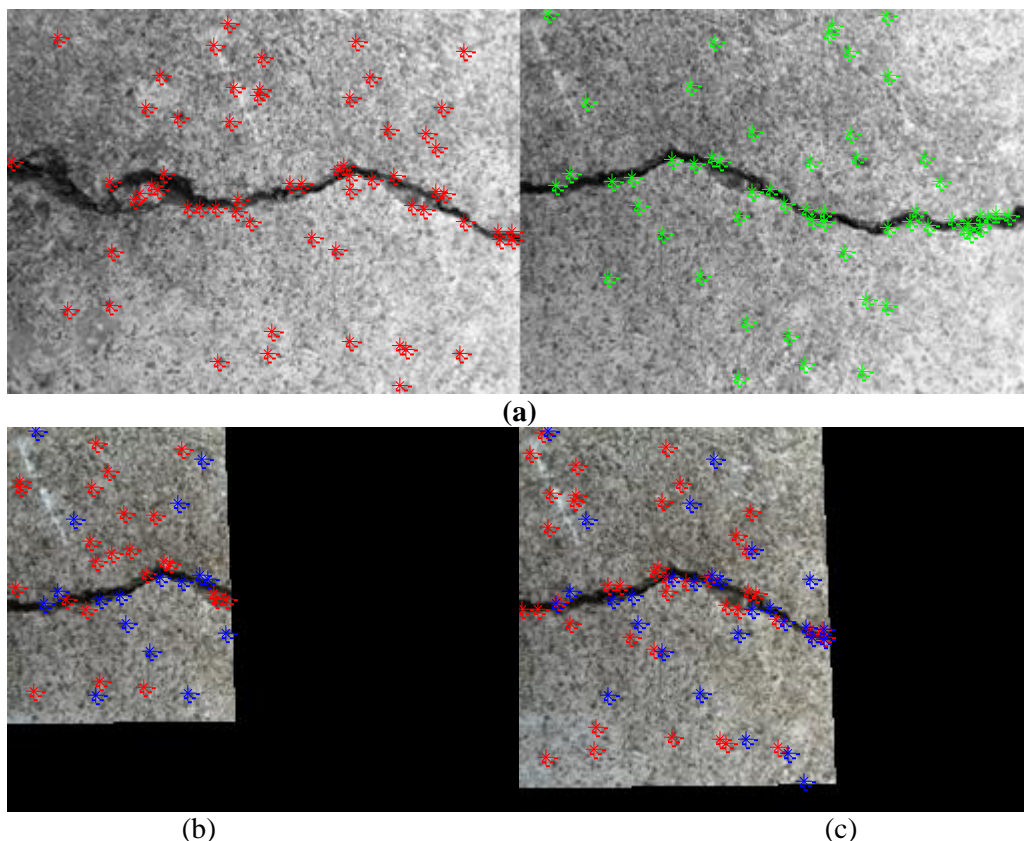


Figure 5. The Searching Region of the Constraining Point based on the Roughly Matched Result, (a) is the Detected Constraint Points of the Image Pair, (b) is the Wrong Matching Result, and (c) is the Correct Matching Result

Considering of the time-consuming and stability of the total algorithm, we choose the edge cover condition as the constraint of the matching result. On this condition, the edge covered points in one image should cover the edge exactly on its matched image edge. From Figure 2(a), the wrong matched result figure, we may found when the matching result is wrong, the edge cover points are not fit the constraint condition to some extent. Figure 5(b,c) shows the overlapped region of the two matching results M'_1 and M'_2 . Comparing with the two results, we may found, when the matching result was correct, the difference of the local characters such as gradient and the gradient direction is small, and vice versa. For the purpose of avoiding the influence of lighting, we use the normalized difference of all the overlapped constrain points as the evaluating of the matching result.

4.3. The Re-matching and Blending of the Image Pair

After each pairs of the resized images have been matched, we save all the pre-matched homograph matrixes M'_i in the computer. Then, we use each homograph matrix M'_i to segment the rough images into overlapped and un-overlapped regions. The extraction of the exact character points of rough image pair will be done only in the overlapped regions. To deduce the calculation and time-consuming, we search the exact character points only according to the position of pre-matched pixels.

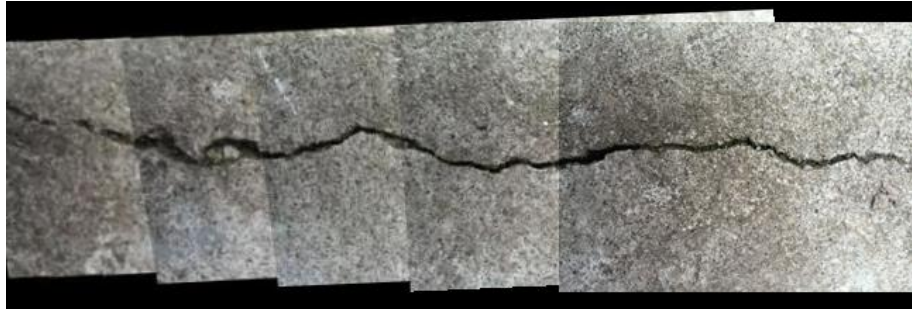


Figure 6. The Matching Result of 5 Sequential Resized Images

5. Experimental Results

To demonstrate the accurate of the matching results, we choose three kinds of image pairs, and calculate these images on Matlab 2010. The three pairs of images of Fig.6 have been experimented in the following experiments. The motivation of our current experiment is to provide the high accurate matching solution to the rough images.



Figure 7. The Blending Result of the Sequential Images



Figure 8. The Local Image of the Matched Sequential Image Results

For the rough images (3264×2448 or larger), we use the above mentioned algorithm to pre-match the resized images of the sequential. Then, we use the calculated matching result, the matching points and homograph matrix M'_i , to realize the accurate matching of the rough image pair. Figure 7 shows the matching result of the first image pair in Figure 6.

To improve the efficient of the rough image pair matching procedure we use the matching points of the resized matching result as the basic searching region. In the following experiments, we choose the 5×5 region centered at the mapped matching pixels as the searching region of the character points. On this condition, the searching region and calculating procedure would be decreased to a large extent. To illustrate the mosaic accurate, we use the partial of the mosaic result centered at overlapped region as the result images. Fig.8 shows the results of three pair images.

6. Experimental Results

Comparing to the other algorithms which aimed at finding out the relative characters of the mosaicking images, our method is more stable, accurate, and adaptable to different kinds of images. On the other hand, all the self-constraint conditions are unstable when it meets the perspective transform. In our algorithm, the constraint condition was adopted after the pre-matching procedure. It improved the stability and made the constraint condition more simply.

In the character point extracting procedure, the distribution and number of character points is the main influence of image matching accurate and stability. In this paper, we extract the character points from strong to weak response and at the same time limit the maximum point number. This algorithm may also improve the stability of the matching processing. Because of the time-consuming is very sensitive to the algorithms, in the proposed algorithm, we avoided all the complicated calculating of the large image, it makes the total time-consuming much less than the other mosaic algorithms.

Acknowledgements

The authors thank the anonymous reviewers for their many valuable comments and suggestions that helped to improve both the technical content and the presentation quality of this paper.

References

- [1] J. Chen, H. Feng, K. Pan, Z. Xu, and Q. Li, "An optimization method for registration and mosaicking of remote sensing images," *Optik - International Journal for Light and Electron Optics*, vol. 125, (2014), pp. 697-703.
- [2] H. Wang, X. Liu, S. Zhang, and W. Quan, "Novel Image Mosaic Algorithm for Concrete Pavement Surface Image Reconstruction," *Procedia - Social and Behavioral Sciences*, vol. 96, (2013), pp. 2692-2697.
- [3] C.-H. Lee, Y.-C. Chiu, T.-T. Wang, and T.-H. Huang, "Application and validation of simple image-mosaic technology for interpreting cracks on tunnel lining," *Tunnelling and Underground Space Technology*, vol. 34, (2013), pp. 61-72.
- [4] Z. Falomir, L. Museros, L. Gonzalez-Abril, and F. Velasco, "Measures of similarity between qualitative descriptions of shape, color and size applied to mosaic assembling," *Journal of Visual Communication and Image Representation*, vol. 24, (2013), pp. 388-396.
- [5] X. Yin and N. Ma, "Study on the automatic Unmanned Aerial Vehicle Image mosaic algorithm," presented at the Intelligent Control and Information Processing (ICICIP), 2011 2nd International Conference on 2011.
- [6] J.-W. Hsieh, "Fast stitching algorithm for moving object detection and mosaic construction," *Image and Vision Computing*, vol. 22, (2004), pp. 291-306.

- [7] K. Chaayasarn, T.-K. Kim, F. Viola, R. Cipolla, and K. Soga, "IMAGE MOSAICING VIA QUADRIC SURFACE ESTIMATION WITH PRIORS FOR TUNNEL INSPECTION," presented at the Image Processing (ICIP), 2009 16th IEEE International Conference on 2009.
- [8] G. Xu and J. Cai, "The Research of Granular Computing Applied in Image Mosaic," presented at the Informatics in Control, Automation and Robotics, 2009. CAR '09. International Asia Conference on (2009).
- [9] A. Elibol, N. Gracias, and R. Garcia, "Fast topology estimation for image mosaicing using adaptive information thresholding," *Robotics and Autonomous Systems*, (2013), pp. 125-136.
- [10] J. Dou and J. Li, "Image matching based local Delaunay triangulation and affine invariant geometric constraint," *Optik - International Journal for Light and Electron Optics*, vol. 125, (2014), pp. 526-531.
- [11] Y. Zou, F. Dong, B. Lei, L. Fang, and S. Sun, "Image thresholding based on template matching with arctangent Hausdorff distance measure," *Optics and Lasers in Engineering*, vol. 51, (2013), pp. 600-609.
- [12] S. Yahiaoui, M. Haddad, B. Effantin, and H. Kheddouci, "Coloring based approach for matching unrooted and/or unordered trees," *Pattern Recognition Letters*, vol. 34, (2013), pp. 686-695.
- [13] U. Kang, M. Hebert, and S. Park, "Fast and scalable approximate spectral graph matching for correspondence problems," *Information Sciences*, vol. 220, (2013), pp. 306-318.
- [14] C. Je and H.-M. Park, "Optimized hierarchical block matching for fast and accurate image registration," *Signal Processing: Image Communication*, vol. 28, (2013), pp. 779-791.
- [15] G. Yongfang, Y. Ming, and S. Yicai, "Study on an Improved Robust Algorithm for Feature Point Matching," *Physics Procedia*, vol. 33, (2012), pp. 1810-1816.
- [16] L. Cheng, M. Li, Y. Liu, W. Cai, Y. Chen, and K. Yang, "Remote sensing image matching by integrating affine invariant feature extraction and RANSAC," *Computers & Electrical Engineering*, vol. 38, (2012), pp. 1023-1032.
- [17] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, (2004), pp. 91-110.
- [18] C. Harris and M. Stephens, "A combined corner and edge detector," presented at the In Proc. of Fourth Alvey Vision Conference (1988).
- [19] L. Fang and X. Wang, "Image segmentation framework using EdgeFlow-Based active contours", *Optik - International Journal for Light and Electron Optics*, vol. 124, (2013), pp. 3739-3745.

Authors



Bo Cai, he received his B.S. degree in mechatronics from China Agricultural University, in 1998, and M.S degree in Beijing University of Posts and Telecommunications, in 2003. Currently, he is the Ph.D. candidate of China Academy of Engineering Physics. His research interests include image processing and detection.



Zhigui Liu, he received the Ph.D. degree in traffic information engineering and control from Southwest Jiao Tong University, China, in 2006. Currently, he is the teacher of Southwest University of Science & Technology, and Ph. D tutor of China Academy of Engineering Physics. His research interests include vision sensor technology, image processing and machine vision, etc.



Junbo Wang, he received his B.S. degree in physics from Sichuan Normal University, China, in 1981, and the Ph.D degree in optical engineering from Chengdu electronic science and technology University, China, in 1985. Currently, he is the teacher of Southwest University of Science & Technology, and Ph. D tutor of China Academy of Engineering Physics. His research interests include atmospheric optical and detection, etc.



Yuyu Zhu, he received his B.S. degree in automation from Southwest University of Science & Technology, in 2002, and M.S degree in control engineering and control theory, in 2009. Currently, he is the teacher of Southwest University of Science & Technology. His research interests include signal detection and processing, power designing etc.