

Image Mosaic Algorithm of Sequential Images Based on Voronoi

Zhicheng Wang, Qian Xu, Zhiheng Wang, Xiaopeng Guo and Yaxing Yang

College of Electronics and Information Engineering, Tongji University, Shanghai, 201804, China

Abstract—The image sequence of Vehicle panorama camera platform has the characteristics of serious distortion, large parallax and large amount of data, which can lead to invalid montage for the image sequence by adopting the common image mosaicking method based on features. In this paper, Voronoi algorithm is introduced according to the actual characteristics of vehicle camera platform. A steady and fast image mosaic method for image sequence is proposed. Firstly, the relative positions between images are structured by utilizing GPS information of images. The splicing region of images are classified through Voronoi algorithm and the confidence is introduced to select the best candidate image; then SIFT feature of image is extracted and image registration is carried out by adopting Moving DLT algorithm, what follows is global alignment for stitching images by using improved Bundle Adjustment algorithm. Finally, the stitching images are integrated by using Multi-Band method. Contrasted with experimental results, the proposed method not only realizes the image mosaic under the condition of large scale, multi parallax and distortion images which can make up for the shortcomings of traditional image mosaicking methods, but also reaches good stitching effect and high mosaicking efficiency.

Keywords—component; image mosaicking; moving DLT; bundle adjustment; multi-band blending

I. INTRODUCTION

Image mosaic is a kind of image registration, splicing, and fusion based on the information of overlapping areas among multiple images. And then get a panoramic image that contain each image information [1]. The technology of image stitching has been used in computer vision, UAV aerial image processing[2], 3-D panoramic roaming[3], advanced driving support system (ADAS) [4] and street maps[2].

The main image mosaic algorithms are based on the overlapping points of the feature points. David G. Lowe proposed the scale invariant feature extraction algorithm SIFT (Scale Invariant Feature Transform)[5] in 1999. Lowe improved the SIFT algorithm and proposed a complete SIFT algorithm[6] in 2004. SIFT features are based on some of the local appearance of the points of interest and regardless of the size and rotation of the image. The tolerance of light, noise, micro-viewing angle is quite high. Matthew Brown and David G. Lowe proposed Auto Stitch[7] that is an automatic image stitching technique and based on the study of SIFT in 2007. The method can splice multiple unordered images of multiple targets and get the corresponding panoramic image. It has become an important milestone in the field of image mosaic, and set off the research boom of the image mosaic technology. Because of the Auto Stitch algorithm uses the global homography matrix to calculate the transformation relationship among images. So it can generate

the issues of ghosting, distortion, and for large parallax image mosaic is helpless. Julio Zaragoza Proposed an APAP (As Projective As Possible) algorithm based on grid transform [8] in 2011. Firstly, the direct linear transformation (DLT) is used to estimate the coherence matrix between two images. the image is divided into multiple dense grids, and the homogeneity relation between each mesh that is calculated and enhance the effect of splicing. Che-Han Chang et al. Proposed a shape-preserving image mosaic method SPHP (Shape-Preserving Half-Projective)[9] in 2014. The method sets different transformation matrices for overlapping and non-overlapping regions and adds similarity constraints to the whole image. Then it corrects the shape of the spliced image and reduces the projection distortion further. In recent years, large parallax images have also made progress. Fan Zhang et al. borrowed the method of classical video debounce and the optimization items and suture splicing method of Content-preserving warps (CPW) in 2014. It improved the splicing performance of large parallax scenes dramatically[10]. Lin, Kaimo et al. improved the suture algorithm, contour detection and straight line detection on the basis of the Parallax-tolerant Image Stitching algorithm in 2016. It added the curve and the structure of the linear structure to keep the structure of the constraint, so that the effect of splicing improved significantly[11].

With the technology of deep learning developing rapidly, some scholars have begun to try to take the deep learning into the image mosaic. Kwang Moo Yi et al. proposed a method of character point detection, direction assignment and description generation by CNN. The method named LIFT (Learned Invariant Feature Transform)[12], in 2016. Daniel DeTone et al. proposed HomographyNet[13] which uses a 10-layer VGG network to regression and predicted the eight parameters of the homograph matrix.

The Voronoi algorithm is often used in the geometric division problem. With the development of computer technology, the scope of Voronoi diagram is expanding gradually[14]. Kise et al. proposed a method of text cutting by the Voronoi diagram, and achieved remarkable results[15]. Oliva et al. optimized the 3D surface reconstruction of irregular objects by the Voronoi algorithm[16].

Image mosaic technology has been applied to the camera platform of the car in the panoramic image mosaic successfully, such as street maps, auxiliary parking system and so on. The application is mainly used in camera platform of the car in a number of cameras to capture the image of the splicing, and then can only get a single point of the panoramic image. But in the camera platform of the car sequence image mosaic. Due to error transmission, image distortion, large amount of data and other

reasons, the common image stitching method is difficult to achieve better splicing. In view of the above problems, this paper presents an effective method for fast splicing of sequence images. Firstly, based on the traditional camera platform of the car, the single point panorama image has been stitched to transform, and get a single point of view. Then the Voronoi algorithm is used to mosaic the sequence image, and the panoramic view of the vehicle trajectory is obtained. According to the experimental results, the method of this paper can be used to stitch the sequence image with large distortion and large data, which is better than the other methods.

II. DATA AND PREPROCESSING

The data sets used in the article are derived from Street View vehicle camera platform. There are three fish-eye cameras on the top of the car and three fisheye images will be obtained in one shot. A panoramic image can be obtained by stitching the image with a predefined algorithm. Casting the panorama image into the sphere coordinate system is the result we saw in the street view map. These panoramic images will be projected onto the sphere coordinate system and provided to the users of the street view map. In order to gain sequential top view images of the roads from these panorama images, we transform the pixels on the lower hemisphere into the plane coordinate system. The image obtained by this method can keep the information of the road as much as possible, however, it will cause a large amount of distortion on object above the road inevitably.

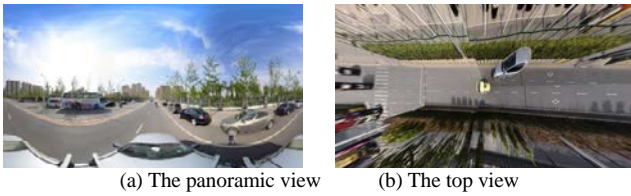


FIGURE I. (A) IS THE PANORAMIC VIEW STITCHED BY THREE FISH-EYE IMAGES, (B) IS THE TOP VIEW TRANSFORMED BY (A).

III. VORONOI GRAPH

The first step in the sequential images stitching algorithm is to extract the critical images with a more or less uniform spatial distribution from the candidate sequences. In the candidate Images, the farther the distance away from the center of the image is, the more serious the object distortion is. Serious distortion will lead to ghosting, blurring and other problems in the mosaic results. Therefore, the next step is to calculate the appropriate area for each image to participate in the splicing.

In mathematics, a Voronoi diagram is a partitioning of a plane into regions based on distance to points in a specific subset of the plane. The Voronoi diagram is a basic geometric structure and commonly used to describe the spatial geometric proximity. We abstract the images into points in the Descartes coordinate system according to the GPS, denoted as $\{P_0, P_1, \dots, P_n\}$.

$$R_k = \{x \in X \mid d(x, P_k) < d(x, P_j)\} \quad (1)$$

$$j = \{0, 1, 2, \dots, n\}, j \neq k$$

R_k is the Voronoi area of the point P_k .

Since the Voronoi diagram can reasonably preserve the effective region of each image, we use the Voronoi method to determine the pixel region of the single image involved in splicing. Firstly, the Voronoi polygon is established based on the position information of the track point after correction, and the image is cropped by the mapping relationship between the track point and the image, and the single image is determined to participate in the splicing area. And the initial image of the initial registration. In order to make the image between the cut after the overlap between the area before the cut in the image area should also be extrapolated.

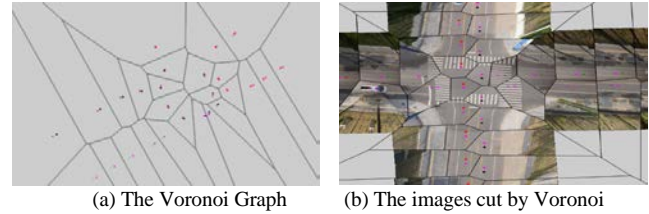


FIGURE II. USE VORONOI TO CHOOSE THE CANDIDATE IMAGES AND COMPUTE THE REGION OF EACH IMAGE

IV. IMAGE ALIGNMENT

In order to further enhance the effect of stitching on the basis of Voronoi., Moving DLT and Bundle Adjustment were used in images alignment. We first extract and match SURF features between all the images, and then use K-D tree to match the feature points by the Euclidean distance, RANSAC is a robust estimation procedure and we use it to estimate the Image transform parameters that has the best consensus with the data. In the case of our algorithm, we select $r = 4$ feature correspondence to compute the homograph H between them using the direct linear transformation(DLT).

According to the given point pair $x_i \Leftrightarrow x'_i, x_i$ and x'_i satisfy the homogeneous vector equation $x'_i = Hx_i$, due to these two vector have the same direction, so

$$x'_i \times Hx_i = \begin{bmatrix} y'_i h^{3T} x_i - w'_i h^{2T} x_i \\ w'_i h^{1T} x_i - x'_i h^{3T} x_i \\ x'_i h^{2T} x_i - y'_i h^{1T} x_i \end{bmatrix} \quad (2)$$

h^{jT} is the j th row in the matrix H , $x'_i = (x'_i, y'_i, w'_i)^T$.

$$\begin{bmatrix} 0^T & -w'_i x_i^T & y'_i x_i^T \\ w'_i x_i^T & 0^T & -x'_i x_i^T \\ -y'_i x_i^T & x'_i x_i^T & 0^T \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix} = 0 \quad (3)$$

Since only two rows of (3) are linearly independent, each pair of points gives two equations for the H element, and the third equation is often omitted in solving H , so that (3) becomes:

$$\begin{bmatrix} 0^T & -w'_i x_i^T & y'_i x_i^T \\ w'_i x_i^T & 0^T & -x'_i x_i^T \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix} = 0 \quad (4)$$

This equation is usually expressed as $A_i h = 0$, where A_i is a 2×9 matrix. h is a 9-dimensional vector consisting of elements of the matrix H .

V. REGION SELECTION

Panoramic Image mosaic using the sections of 2-4 gives a preliminary result (see figure). Each image acquisition vehicle in the candidate images will appear in the final image. A large number of cars in the road will block the traffic signs. Consider that each region is covered by multiple Image, we try to choose a most suitable image to cover the current region to avoid too much cars appear in the mosaic results. Let C_0 be the confidence of the current image, we choose $k = 5$ nearest images, note as P_i , and the image used to cover the current region is computed as:

$$\hat{I} = \arg \min_I \|d_i \cos \theta_i\|^2 \quad (5)$$

d_i means the distance between P_i and P_0 , θ_i is the difference between the angle of the cars in the images.



(a) Without region selection



(b) With region selection

FIGURE III. COMPARISON OF REGION SELECTION

VI. EXPERIMENTS

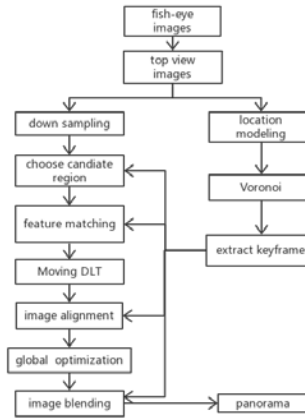
This paper using the experimental data from street map panoramic data collection platform, which carrying three fisheye camera on-board platform for data collection. The data is mainly used for stitching 360 street streetscape. In order to extract the road view from the original dataset, firstly, we project the panoramic image through predefined algorithm. Then we get a top view which consist of 1200 pixels length and 3600 pixels high. Some of the original drawings to be spliced are shown in Figure 4. It can be seen from the image that although a single image contains more scenes and most of the regions have different degrees of distortion. The smaller distortion areas are

concentrated on the road which near the collecting vehicle, and the objects above the road surface will exhibit radial distortion. It can hardly extract the effective feature matching regions.

There are algorithm steps used in this experiment of this paper. The main idea of these algorithms is to use the Voronoi algorithm to model all the frames, extract the key points, partition the region and select the candidate images, and using the image matching algorithm for local matching. The Bundle Adjustment is used for Global alignment. At last, the fusion algorithm is used to fuse the matching results.



(a) part of the data



(b) The implemented stitching pipeline

FIGURE IV. EXAMPLES OF THE CANDIATE IMAGES AND THE IMPLEMENT STITCHING PIPELINE



(a) One way street



(b) Dual carriageway



(c) Crossroad

FIGURE V. EXAMPLES OF TOP VIEW PANORAMA STITCHED BY OUR METHOD

A. Quantitative Analysis

MSSIM algorithm is an evaluation index of full-reference image quality, the image quality is evaluated from three aspects. That is brightness, contrast, structure respectively. The value range of SSIM is $[0, 1]$, when closer to 1 the value is, higher the similarity of the two images is, lower the distortion of the target image is. The target image and the reference image are provided, the similarity between the two images is:

$$SSIM(x, y) = \frac{(2u_x u_y + c_1) (2\sigma_{xy} + c_2)}{(u_x^2 + u_y^2 + c_1) (\sigma_x^2 + \sigma_y^2 + c_2)} \quad (6)$$

Where u_x, u_y represent the average of x, y respectively.

σ_x^2, σ_y^2 represent variance of x and y . σ_{xy} is covariance of the two images. c_1, c_2 are the constant. In order to avoid the circumstance that the value of Denominator is 0:

$$c_1 = (k_1 L)^2, c_2 = (k_2 L)^2 \quad (7)$$

Where: $K1=0.01, K2=0.03, L=255$. In this experiment, because there is no ideal reference image, the random sampling method is adopted to select multiple rectangular regions from the mosaic result to value.

The total number of rectangles selected is N , given the circumstance that the Rectangular region has an effect on the result, The average, variance and covariance of each rectangle region are calculated by using Gauss weighting, then the SSIM of corresponding region is calculated. Finally, the average value is taken as the measurement standard on structural similarity of two image, that is, the average structure similarity MSSIM.

TABLE I. RESULT OF MSSIM

	Without Voronoi	With Voronoi
Street	0.517	0.645
Expressway	0.583	0.686
Sideway	0.559	0.643

VII. CONCLUSION

In this paper, a fast mosaic method of sequence images based on Voronoi is proposed. This method covers the shortage of the traditional algorithm in the large distortion image stitching, and solves the problem of the vehicle sequence panoramic image stitching. Experiments show that the method is stable and fast and has good mosaic effect on single road, dual carriageway, crossroads and other sections. Real world application also proves the effectiveness of proposed method. In the future, we can improve the algorithm from single panoramic image to top view image. In this way, the distortion area of the single point top view is reduced as much as possible, so as to improve the final stitching effect.

REFERENCES

- [1] Szeliski, Richard. "Image alignment and stitching: A tutorial." *Foundations and Trends® in Computer Graphics and Vision* 2.1 (2006): 1-104.
- [2] Eschmann, Christian, et al. "Unmanned aircraft systems for remote building inspection and monitoring." *Proceedings of the 6th European Workshop on Structural Health Monitoring, Dresden, Germany*. Vol. 36. 2012.
- [3] Adel, Ebtsam, Mohammed Elmogy, and Hazem Elbakry. "Image stitching based on feature extraction techniques: a survey." *International Journal of Computer Applications* (0975-8887) Volume (2014).
- [4] Liu, Yu-Chih, Kai-Ying Lin, and Yong-Sheng Chen. "Bird's-eye view vision system for vehicle surrounding monitoring." *Robot vision* (2008): 207-218.
- [5] Lowe, David G. "Object recognition from local scale-invariant features." *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*. Vol. 2. Ieee, 1999.
- [6] Lowe, David G. "Distinctive image features from scale-invariant keypoints." *International journal of computer vision* 60.2 (2004): 91-110.
- [7] Brown, Matthew, and David G. Lowe. "Automatic panoramic image stitching using invariant features." *International journal of computer vision* 74.1 (2007): 59-73.
- [8] Zaragoza, Julio, et al. "As-projective-as-possible image stitching with moving DLT." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2013.
- [9] Chang, Che-Han, Yoichi Sato, and Yung-Yu Chuang. "Shape-preserving half-projective warps for image stitching." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
- [10] Zhang, Fan, and Feng Liu. "Parallax-tolerant image stitching." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
- [11] Lin, Kaimo, et al. "Seagull: Seam-guided local alignment for parallax-tolerant image stitching." *European Conference on Computer Vision*. Springer International Publishing, 2016.
- [12] Yi K M, Trulls E, Lepetit V, et al. LIFT: Learned Invariant Feature Transform[M]// *Computer Vision – ECCV 2016*. Springer International Publishing, 2016.
- [13] DeTone, Daniel, Tomasz Malisiewicz, and Andrew Rabinovich. "Deep image homography estimation." *arXiv preprint arXiv:1606.03798* (2016).
- [14] Aurenhammer, Franz. "Voronoi diagrams—a survey of a fundamental geometric data structure." *ACM Computing Surveys (CSUR)* 23.3 (1991): 345-405.
- [15] Kise, Koichi, Akinori Sato, and Motoi Iwata. "Segmentation of page images using the area Voronoi diagram." *Computer Vision and Image Understanding* 70.3 (1998): 370-382.
- [16] Oliva, J - M., M. Perrin, and Sabine Coquillart. "3D reconstruction of complex polyhedral shapes from contours using a simplified generalized Voronoi diagram." *Computer Graphics Forum*. Vol. 15. No. 3. Blackwell Science Ltd, 1996.