Smart Structures and Systems, *Vol. 25, No. 5 (2020) 631-641*
DOI: https://doi.org/10.12989/sss.2020.25.5.631

Vision-based full-field panorama generation by UAV using GPS data and feature points filtering

Yapeng Guo [1], Yang Xu [2], Haowei Niu [1], Zhonglong Li [1], Yuhui E. [3], Xinghua Jiao [3] and Shunlong Li*[1]

[1] *School of Transportation Science and Engineering, Harbin Institute of Technology, 73 Huanghe Road, Harbin 150090, China*
[2] *School of Civil Engineering, Harbin Institute of Technology, 73 Huanghe Road, Harbin 150090, China*
[3] *Liaoning Transportation Development Center, 128 Shashan Road, Shenyang 110000, China*

(*Received September 19, 2019, Revised November 15, 2019, Accepted December 5, 2019*)

Abstract. To meet the urgent requirements of safety surveillance from civil engineering management authorities, this study proposes a refined and efficient approach to generate full-field high-resolution panorama of construction sites using camera-amounted UAV (Unmanned Aerial Vehicle). GPS (Global Position System) information extraction for pre-registration, feature points filtering for efficient registration and optimal seaming line seeking for fusion are performed in sequence to form the full-field panorama generation framework. Advantages of the proposed method are as follows. First, GPS information can sort images for pre-registration, avoiding inefficient repeated pairwise calculations and matching. Second, the feature points are filtered according to the characteristics of the construction site images to reduce the amount of calculation. The proposed framework is validated on a road construction site and results demonstrate that it can generate an accurate and high-quality full-site panorama for the safety supervision in a much efficient manner.

Keywords: full-field panorama; UAV; GPS information; image registration; image stitching

1. Introduction

For the autonomous and full-field safety surveillance, assessment and management in civil engineering, e.g., structures and roads, camera-mounted UAVs have shown evident advances in many cases (Ellenberg *et al.* 2014, Li *et al.* 2015, Xu *et al.* 2015, Ham *et al.* 2016, Irizarry and Costa 2016, De Melo *et al.* 2017). It enables the efficient collection of image information of civil structures from various perspectives, which is superior to the closed-circuit television (CCTV) cameras (Liu *et al.* 2014). UAV-based inspection and monitoring becomes a safer, more economical and less time-consuming solution compared with other means (Kim and Kim 2018), and has been gradually used by civil engineering management authorities. Recent breakthroughs of computer vision algorithms further help improving UAV-based monitoring and condition assessment in time. Metni and Hamel (2007) described the dynamics of a UAV for the monitoring of structures and maintenance of bridges, and then presented a novel control law for quasi-stationary flights above a planar target based on computer vision. Reagan *et al.* (2016) developed an autonomous UAV in conjunction with the 3D DIC (Digital Image Correlation) technique for monitoring bridges. Hackl *et al.* (2017) investigated the use of a UAV and modern photogrammetric technology to obtain topographical information for the bridge risk assessment. Although advantages and effectiveness of UAV-based inspections and monitoring compared with conventional methods have been demonstrated in the aforementioned cases, the most vital challenge is to find an efficient, prompt and accurate combination manner of enormous visible images acquired by UAVs from varying angles and positions. Full-field monitoring and management based on image stitching techniques provide a possible solution for these issues.

Image stitching techniques are widely adopted during the post-processing of UAV-obtained images to accomplish feature extraction, matching and form wider views of surveilled targets from single frames. Xiang *et al.* (2014) developed a fast mosaic system of massive UAV images using GPS and POS data. To solve the dislocation problem in UAV image stitching, Li *et al.* (2017) used a new energy aggregation and traversal strategy to find a more optimal seam, which overcome the theoretical limitation of classical algorithms. Bang *et al.* (2017) generated a high-quality panorama of a construction site to help managers to easily identify various conditions by image stitching, which consisted of three modules: blur filtering, key frame selection, and camera correction. Zhang *et al.* (2018) introduced an image mosaic technique based on Speed Up Robust Feature (SURF) to solve the rapid splicing problem of UAV low-altitude remote sensing images. Akbar *et al.* (2019) developed an autonomous monitoring system for structural defect detection by UAVs, in which images were stitched together to form the complete view and then damage detection was performed. The traditional image stitching techniques are faced with disadvantages of low efficiency and limited speed in the post-processing, blurring and ghost effects in the fusion image.

∗Corresponding author, Professor,
E-mail: lishunlong@hit.edu.cn

ISSN: 1738-1584 (Print), 1738-1991 (Online)

This study aims to provide an entire high-resolution full-field panorama generation methodology, simultaneously holding advantages of making full use of UAV's precise GPS information for pre-registration, filtering feature points to accelerate and avoiding blurring and ghost effects with optimal seaming line. The remainder of this paper is arranged as follows. Section 2 illustrates the methodology of proposed panorama generation framework based on refined GPS information extraction, feature points filtering for efficient registration and image fusion by optimal seaming line seeking. Section 3 describes the implementation details of the validation of the proposed approach on a highway road construction site. Section 4 concludes the paper.

2. Full-field panorama generation methodology

Fig. 1 describes the overall panorama generation framework for the full-site surveillance of civil engineering utilizing high-resolution UAV images in this study, which includes a sequence of procedures containing: (I) GPS information extraction from UAV images and coordinates transformation from the geographic coordinate system to local plane coordinates; (II) feature region extraction and feature points filtering for Homography matrix computation; (III) image fusion based on optimal seaming line seeking to eliminate the blurring and ghost effects.

2.1 Image pre-registration using GPS information

A UAV integrated with Inertial Navigation Unit (INU) and Global Position System (GPS) can generate a 6-element-wise tuple $O_i(\psi_i, \theta_i, \varphi_i, x_i, y_i, z_i)$ at every photography location $\boldsymbol{O}_i$, in which $\psi, \theta, \varphi, x, y, z$ represent roll, pitch, course angles and spatial coordinates in the *x*-, *y*- and *z*-directions, respectively. Gauss-Krüger projection (Carlos 2008) could be used to transform the geographic coordinate system (L, B) to the plane coordinates (x, y) as

$$
\begin{aligned}
x &= s + \frac{L^2 N}{2} \sin B \cos B \\
&\quad + \frac{L^4 N}{24} \sin B \cos^3 B \,(5 - \tan^2 B + 9\eta^2 + 4\eta^4) + \cdots \\
y &= LN \cos B + \frac{L^3 N}{6} \cos^3 B \,(1 - \tan^2 B + \eta^2) \\
&\quad + \frac{L^5 N}{120} \cos^5 B \,(5 - 18 \tan^2 B + \tan^4 B) + \ldots
\end{aligned}
\tag{1}
$$

where L, B denotes the longitude and latitude in the geographic coordinate system, respectively; s denotes the meridian arc length from equator to latitude B; N denotes the radius of curvature in prime vertical, η is an intermediate variable. They are determined by the earth geometric parameters as

$$
\begin{aligned}
& a = 6378137, && f = \frac{1}{298.257223563}, \\
& b = a(1 - f) && \\
& e = \frac{\sqrt{a^2 - b^2}}{a}, && e' = \frac{\sqrt{a^2 - b^2}}{b}, \\
& N = \frac{a}{\sqrt{1 - e^2 \sin^2 B}}, && \eta = e' \cos B
\end{aligned}
\tag{2}
$$

where a, b, f, e, e' denote the half length of major, minor axis, flattening, first and second eccentricity ratio of the

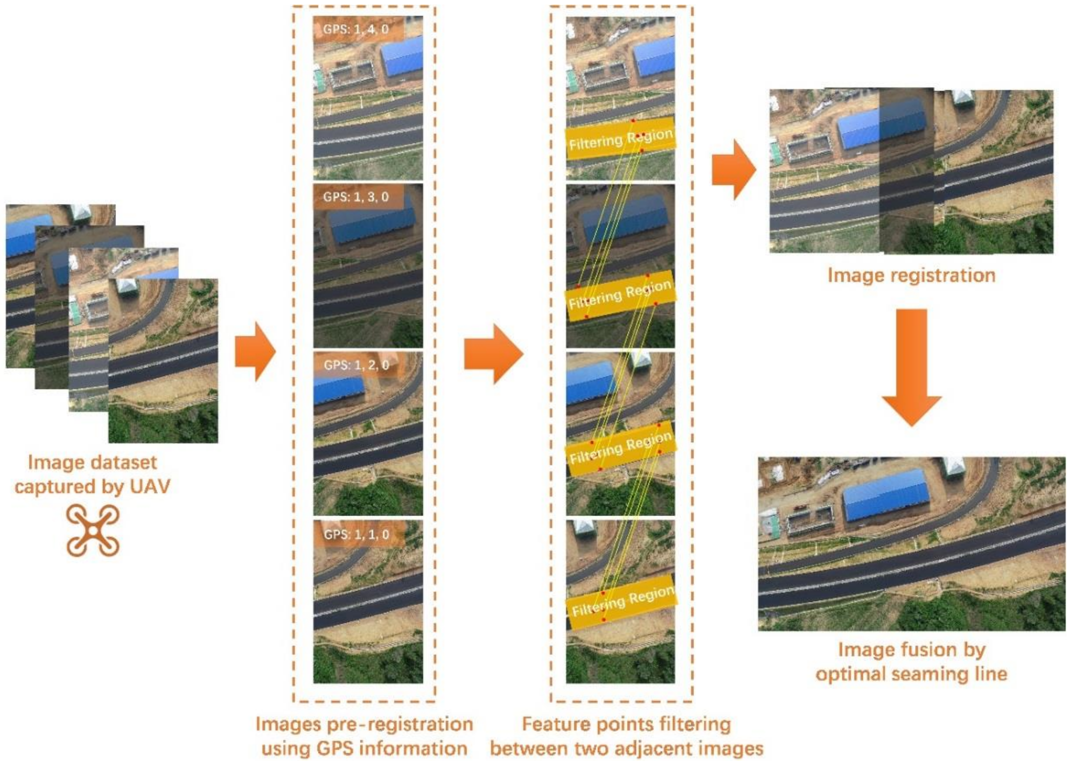


Fig. 1 Overall schematic of the proposed panorama generation framework

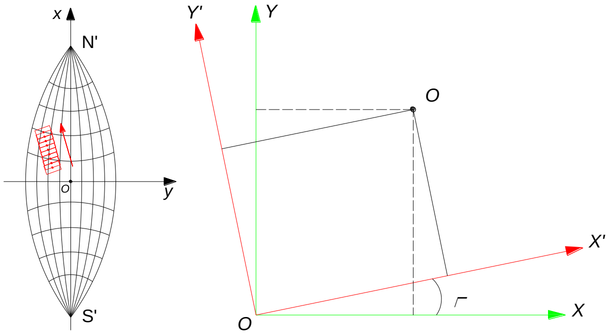


Fig. 2 Local plane coordinate transformation

earth ellipse, respectively.

Always the UAV route in the Cartesian coordinate system is oblique straight line and the aerial images are also tilted, which prevent the convenience of stitching calculation. Thus, in this study UAV-captured images would be arranged in order along Y or X axis by two-dimensional coordinate transformation, which mainly lies in the changes of course angle θ, as shown in Fig. 2. The rectification process of coordinate rotation can be expressed as

$$\begin{aligned} x' &= x \times cos(\theta) + y \times sin(\theta) \\ y' &= y \times cos(\theta) - x \times sin(\theta) \end{aligned} \tag{3}$$

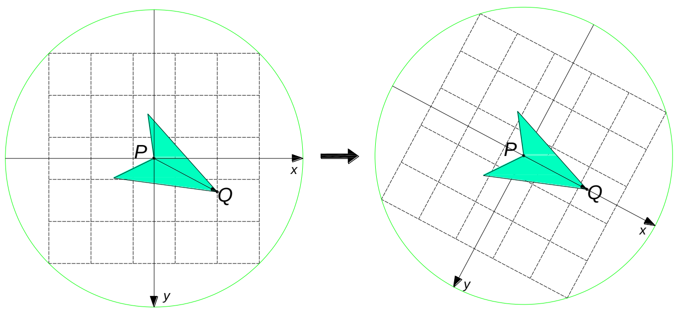
where *x*, *y* denotes the original coordinates and *x'*, *y'* denote the rectified coordinates considering rotation effects, respectively.

According to the coordinates of images, only (n-1) matches need to be finished instead of n(n-1)/2 before pre-registration.

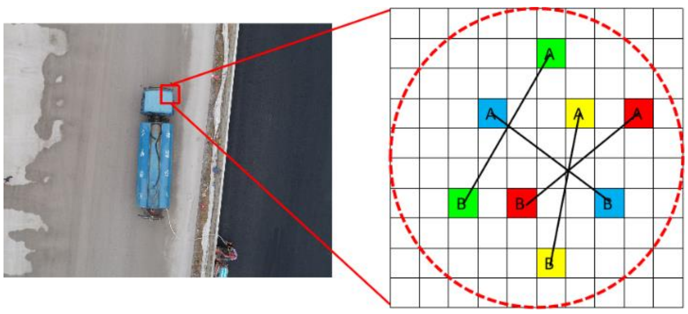
2.2 Feature points filtering for efficient registration

Most existing multi-image registration techniques compute the transformation matrix of images based on feature points. As the image resolution is getting higher and higher, conventional methods for feature point selection, e.g., Scale Invariant Feature Transform-SIFT (Matthew and Lowe 2007), need to filtrate all the pixels in the overlap range, leading to enormous calculation costs. The civil engineering image often presents more structural features due to the particularity of the construction, which provides feasibility for the extraction of the feature area. These feature regions usually include: roads, foundation pits, bridge columns and so on. After extracting the feature

(a) FAST corner point detection



(b) Establishing process of new local coordinate system



(c) BRIEF feature descriptor

Fig. 3 ORB feature point extraction and description regions, there is no need to detect feature points in the whole image, instead only in the feature regions.

Many feature points have been explored to describe specific features. The ORB feature point (Rublee *et al.* 2011) are widely used and has been proved for better performance. Thus, the ORB feature point would be employed in this study. ORB (Oriented FAST and Rotated BRIEF) is an advanced feature extraction mode and developed on the basis of FAST (Features from Accelerated Segment Test) corner detector (Rosten and Drummond 2006) and BRIEF (Binary Robust Independent Elementary Features) feature operator (Calonder *et al.* 2010), thus inheriting the advantages of requiring low calculation costs.

The feature points detected by FAST corners are invariant to translation and rotation and have good robustness to noise in the calculation process. As shown in Fig. 3(a), the candidate corner point $\boldsymbol{P}$ would be selected if the differences between its intensity I_p and all the intensity of neighboring points $I_x, x \in \{p(\delta)\}$ yield

$$\begin{aligned} &\text{num}(x) \geq n \\ &s.t. |I_p - I_x| \geq m, x \in \{P(\delta)\} \end{aligned} \tag{4}$$

where n and m are the pre-set thresholds of point numbers and intensity differences, respectively. x represents the individual point belong to the neighborhood of $\boldsymbol{P}$ within a range of δ pixels. In this study, $\delta = 3$, $n = 9$ is the default setup referring to (Rosten and Drummond 2006).

FAST corner detector does not describe the direction of feature points. thus, the intensity moment is employed to determine the direction and uniqueness of feature points in the extraction process. As shown in Fig. 3(b), for all the pixels in the neighborhood $x \in \{P(\delta)\}$ with a center point $\boldsymbol{P}$ and a radius of δ, the intensity centroid and *pq*-order moments of x- and y- axis can be generally expressed as

$$m_{pq} = \sum_{(x,y) \in P(\delta)} x^p y^q \, I(x,y) \tag{5}$$

where m_{pq} represents the *pq*-order intensity moment, $I(x,y)$ represents the gray-scale pixel value at pixel location (*x*, *y*); *p* and *q* denote the orders of x- and y-coordinates, respectively. The location of intensity centroid on the image coordinates is determined by

$$Q = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}\right) \tag{6}$$

Then, a novel local coordinate system is established, whose origin center is point $\boldsymbol{P}$ and positive X-axis direction is the principle direction of vector $\overrightarrow{PQ}$. The corresponding Y-axis direction is automatically determined to be perpendicular to X-axis clockwise in the image coordinates as shown in Fig. 3(b).

Afterwards, the BRIEF feature descriptor is defined and shown in Fig. 3(c). For the local coordinate system is determined by the direction of vector $\overrightarrow{PQ}$, the rotation invariance of feature descriptor $\boldsymbol{D}_{\text{BRIEF}}$ is guaranteed and used to accomplish point matching under different circumstances if

$$\begin{aligned} &\frac{num(\boldsymbol{D}_{\text{BRIEF}} = \widetilde{\boldsymbol{D}}_{\text{BRIEF}})}{T} \geq r, \boldsymbol{D}_{\text{BRIEF}} = \{\tau(A_i, B_i)\} \\ &\tau(A_i, B_i) = \begin{cases} 1, \text{ if } & I(A) < I(B) \\ 0, \text{ otherwise} \end{cases}, i = 1 \sim T \end{aligned} \tag{7}$$

where $\boldsymbol{D}_{\text{BRIEF}}$ and $\widetilde{\boldsymbol{D}}_{\text{BRIEF}}$ are two BRIEF descriptors for two points to be matched under their individual circumstance, $num(\boldsymbol{D}_{\text{BRIEF}} = \widetilde{\boldsymbol{D}}_{\text{BRIEF}})$ represents the number of same elements for $\boldsymbol{D}_{\text{BRIEF}}$ and $\widetilde{\boldsymbol{D}}_{\text{BRIEF}}$; *I*(*A*) and *I*(*B*) represent the pixel intensities of point *A* and *B*, respectively; *T* represents the length of string and generally *T* = 512 (Calonder *et al.* 2010); *r* denotes the pre-set threshold ratio and is set to 0.95 in this study. And then, the Homography matrix $\boldsymbol{H}_{i+1,i}$ would be updated based on the pixel coordinates of feature points.

2.3 Image fusion by optimal seaming line seeking

Assume two overlapped images I_i and I_{i+1} as input, the output should be their fusion image, i.e., $\hat{I}_{i+1,i} = \text{Fusion}(I_i, I_{i+1})$, in which the seaming line plays the most important role. The consideration of optimal seaming line searching is two-fold. In terms of pixel intensity, differences between pixels on the seaming line and in two original images are required to reach as slight as possible, which is defined as color difference; geometrically, pixels on the seaming line are required to have the most similar structures with the original images, which is defined as geometry difference. The optimal seaming principle aims to find a seam line, reaching the minimum of total differences for all the points on it

$$\begin{aligned} &min \sum_{(x,y) \in \Re} E(x,y) \\ &E(x,y) = E_{color}(x,y)^2 + E_{geometry}(x,y) \end{aligned} \tag{8}$$

where $E(x,y)$, $E_{color}(x,y)$, $E_{geometry}(x,y)$ represents the total, color intensity and structural geometry differences of the overlapped pixels, respectively. For every individual point $\boldsymbol{P}$ located in the range $\Re$, $(x,y) \in P_{3\times3}$ represents all the points in its 3×3 neighborhood. For these nine pixels, correlations between the four edge corner pixels and the central pixel are emphasized. The evaluation metric of geometric structure similarity is based on gradient operators in x- and y- directions as

$$S_x = \begin{bmatrix} -2 & 0 & 2 \\ -1 & 0 & 1 \\ -2 & 0 & 2 \end{bmatrix}, \quad S_y = \begin{bmatrix} -2 & -1 & -2 \\ 0 & 0 & 0 \\ 2 & 1 & 2 \end{bmatrix} \tag{9}$$

When S_x or S_y slides over the range $\Re$, the corresponding convolution results are generated accordingly as

$$\begin{aligned} &E_{x,i} = S_x \otimes I_i, E_{y,i} = S_y \otimes I_i \\ &E_{x,i+1} = S_x \otimes I_{i+1}, E_{y,i+1} = S_y \otimes I_{i+1} \end{aligned} \tag{10}$$

where $\otimes$ denotes the convolution operator, summing all the dot multiplication results in the 3×3 neighborhood. The procedures for optimal seaming line searching are shown in Fig. 4.

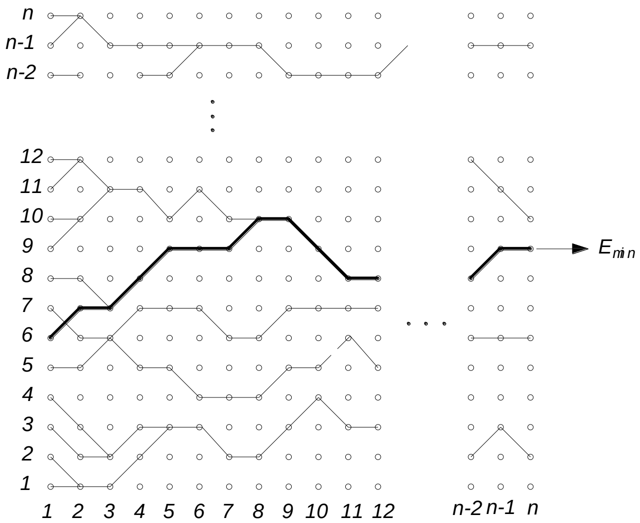


Fig. 4 Procedures for optimal seaming line searching

Afterwards, edge fusion is performed to eliminate ghost and blurring effects on the input original images by a weighted sum of overlapped original images as

$$I(x,y)=\sum_{i=1}^{2} d_i(x,y)I_i(x,y)\,,(x,y)\in\Re \tag{11}$$

where (x,y) represents the pixel position in the image coordinates in the range $\Re$, which has been previously selected from the overlapped region of stitched images I_1 and I_2; $I(\cdot)$ and $I_i(\cdot)$ represents the pixel intensity after fusion and the *i* th original image intensity respectively; d_i denotes the weighted coefficients, and is set ranging from 0 to 1 by

$$d_i(x,y)=\frac{1}{h}y, 0\le y\le h \tag{12}$$

where *h* represents the height of the overlapping region.

3. Case study: Express-way construction site

The investigated road construction site is part of provincial middle ring express-way with a total route length of 118.978 km, a design speed of 80 km/h and a roadbed width of 24.5 meters. The surveilled construction section has a length of about 120 meters and width of about 50 meters. Fig. 5 shows the investigated road construction site and DJI Phantom4Pro UAV employed in this study. The satellite positioning module of DJI Phantom4Pro UAV is dual modes of GPS and GLONASS; visual positioning and obstacle perception systems are integrated for perceptions. The vertical and horizontal hover accuracies are 0.1~0.5 m and 0.3 m~1.5 m, respectively. The maximum horizontal

Fig. 5 Surveilled road construction site and used UAV

Fig. 6 Representative images of the surveilled road construction site

flight and rotation speeds are 72 km/h and 250°/s. The imaging sensor is a CMOS with a size of 2.54 cm × 2.54 cm, equipped with a maximum resolution of 5472 × 3684. The FOV angle of the camera lens is 84°, focal length is 8.8 mm/24 mm, and diaphragm is f/2.8 to f/11. The rotating range of platform holder for stabilization is -90° to 30° with a control accuracy of 0.03°. The battery has a capacity of 5870 mAh and the flight time is about 30 min. The fps of the UAV camera is 30. The cruising control system is PIX4D.

Fig. 6 shows some representative images of the surveilled road construction site, which are continuously shot at a constant height of 30 meters and with an overlap ratio of 50%, containing a variety of on-site objects, e.g., trucks and workers in the asphalt paving. The resolution of captured original images is 3684×5472. The captured sequential images are taken along the main road direction with obvious linear features. Therefore, the following stitching process can be mainly performed along the road direction.

3.1 Image pre-registration using GPS information

According to the camera pose information recorded by the airborne GPS, INU and altimeter integrated in the UAV, the following parameters of the aerial image are extracted including longitude, latitude, altitude, pitch angle, course angle and roll angle. Table 1 shows the corresponding extracted GPS-POS parameters for the representative images in Fig. 6.

It can be seen from Table 1 that when the UAV is photographing during flight, the cruising height is approximately constant, pitch and roll angles are always stable benefiting from the self-balancing system. Translational and course angle changes in the cruising plane are emerging all the time, which will cause fluctuation to the image stitching results. The possible reason is that the wind influence is inevitable.

To eliminate the fluctuation disturbances in the cruising plane caused by longitude and latitude translations, position rectification is performed based on the Gauss-Kruger projection theorem, which has been illustrated in Section 2.1. Table 2 shows the position coordinate transformation results based on GPS information and Gauss-Kruger projection. Afterwards, local coordinates are transformed and generalized to eliminate the influences caused by course angle changes and ensure that all the captured images are arranged in a global coordinate system. Table 2 also shows generalized local coordinate transformation results of latter images to the first one. According to the results, the above images can be sorted as shown in Fig. 7. Only 8 matches need to be done instead of 36 before pre-registration.

Table 1 Extracted GPS-POS parameters for representative original images

DJI Image	Longitude (E)	Latitude (N)	Altitude (m)	Pitch angle θ (°)	Course angle φ (°)	Roll angle ψ (°)
1	124.0559178	41.91028614	173.449	-90	-18.2	0
2	124.0555827	41.91099425	173.649	-90	-19.3	0
3	124.0558695	41.91038542	173.449	-89.9	-18.9	0
4	124.0556302	41.91089447	173.649	-90	-18.7	0
5	124.0558174	41.91049372	173.449	-89.9	-19.1	0
6	124.0556769	41.91079408	173.549	-89.9	-18.8	0
7	124.0557694	41.91059458	173.449	-90	-19.1	0
8	124.0557247	41.91069328	173.549	-90	-19.1	0
9	124.0555372	41.91108517	173.649	-90	-18.4	0

Table 2 Gauss-Kruger projection plane coordinates and local coordinates transformation

DJI Image	Position coordinates		Local coordinates		Local coordinate transformation	
	x	y	x	y	x'	y'
1	587606.9097	4642211.399	0	0	0	0
2	587578.1473	4642289.704	-28.76	78.3	0.71	83.42
3	587602.7636	4642222.396	-4.15	11	0	11.75
4	587582.2246	4642278.676	-24.69	67.28	0.64	71.66
5	587598.3061	4642234.345	-8.6	22.95	0.04	24.51
6	587586.2335	4642267.585	-20.68	56.19	0.48	59.87
7	587594.1811	4642245.496	-12.73	34.1	0.12	36.39
8	587590.3354	4642256.434	-16.57	45.03	0.38	47.99
9	587574.2431	4642299.777	-32.67	88.38	0.61	94.22

Fig. 7 Pre-registration images of the surveilled road construction site

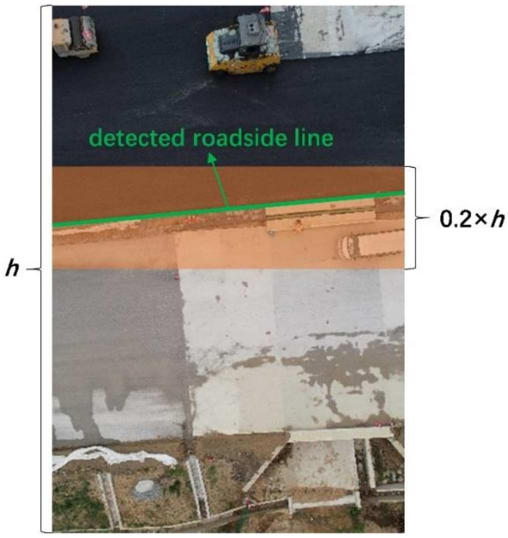


Fig. 8 Definition of the filtering feature region

3.2 Feature points filtering for efficient registration

In this case study, images were acquired at expressway construction site. Striped roads can be used as features for filtering region extraction. A clear road edge can be observed in the image, so this study considers employing Hough line detection method to extract the roadside line. As shown in Fig. 8, the feature region is defined as an area whose center is the center of the detected roadside line and height is 20% of the image's.

After extracting the feature region, ORB feature points are detected in these regions. In Fig. 9, left is feature points detection in the whole image and right is in the feature region. The number of points is significantly reduced. In Table 3, it indicates that there is an average of 13.44% reduction of computation consumption. Then, Homography matrix is computed based on matching feature points. The images are stitched according to Homography matrixes as shown in Fig. 10.

Fig. 9 Feature points detection results: left is in the whole image and right is in the feature region

3.3 Image fusion by optimal seaming line seeking

In addition to the improvement of the processing speed, the proposed image stitching method generates a clearer

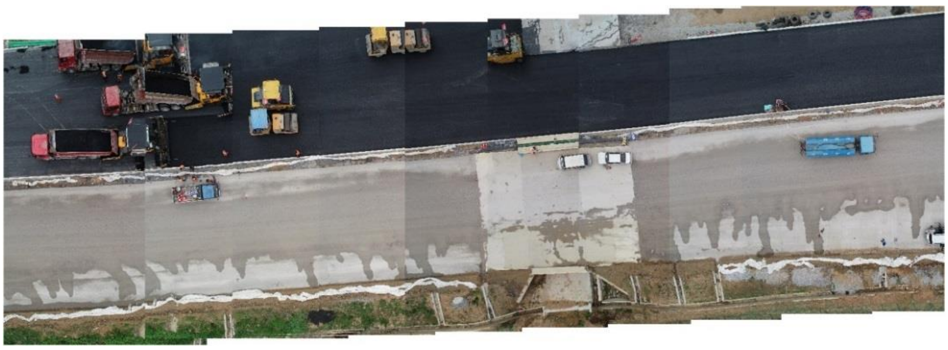


Fig. 10 Image stitching results based on Homography matrixes

Table 3 Homography matrixes and number of feature points using ORB feature points

Homography	Number of feature points			Elements		
	Before	After	Ratio			
$H_{1,3}$	781	85	10.88%	0.51	-0.02	14.85
				-1.10	0.72	339.35
				0.00	0.00	1.00
$H_{3,5}$	842	141	16.75%	1.00	0.00	-6.46
				0.04	1.11	241.12
				0.00	0.00	1.00
$H_{5,7}$	948	106	11.18%	1.04	0.00	-9.20
				0.07	1.03	238.49
				0.00	0.00	1.00
$H_{7,8}$	885	154	17.40%	0.91	0.01	-9.68
				-0.12	0.97	255.84
				0.00	0.00	1.00
$H_{8,6}$	1015	150	14.78%	0.71	-0.03	8.41
				-0.51	0.67	291.94
				0.00	0.00	1.00
$H_{6,4}$	990	131	13.23%	1.21	0.04	-31.78
				0.32	1.30	226.20
				0.00	0.00	1.00
$H_{4,2}$	879	57	6.48%	1.04	-0.03	-8.53
				0.32	1.02	211.71
				0.00	0.00	1.00
$H_{2,9}$	1119	188	16.80%	0.98	0.01	-8.96
				0.07	1.00	172.10
				0.00	0.00	1.00

(a) Optimal seaming line (b) Blurring ghost effect

Fig. 11 Two image stitching comparison between proposed and traditional method

result with optimal seaming line seeking and edge fusion. Based on the proposed ORB features, the optimal seaming line, shown in Fig. 11(a), is sought out with the input of two adjacent images. The black line represents the optimal seaming line, which possesses the minimum of total differences for all the points on it. Without optimal seaming

Fig. 12 Full-site panorama of the entire road construction site by ICE

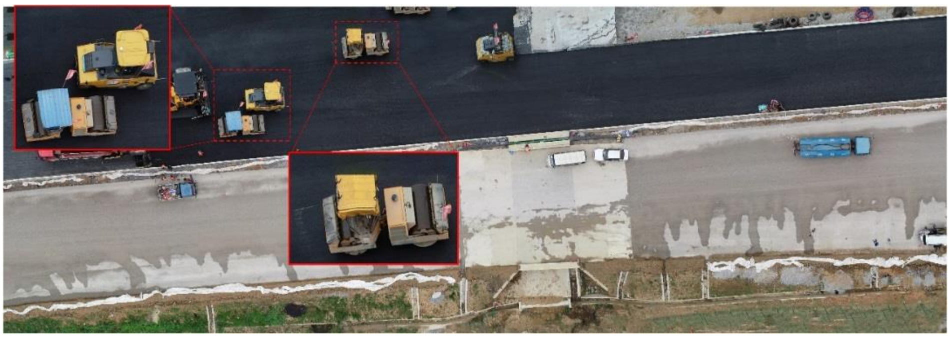


Fig. 13 Full-site panorama of the entire road construction site by the proposed framework

line seeking and edge fusion, blurring, ghost effects are inevitable for the traditional image stitching method, especially for moving objects, which is shown in Fig. 11(b). This comparison demonstrates the significant need for the optimal seaming line seeking and edge fusion of the proposed method.

3.4 Comparison and discussion

Finally, the fast image stitching algorithm is applied to all the images and a full-field panorama of the surveilled construction site is obtained, which is shown in Fig. 13.

The research direction of image stitching focuses on multiple homography matrixes for one image (Lin *et al*. 2015), rather than the conventional method, i.e., one homography matrix for one image. Undoubtedly, this change is able to find the corresponding transition relationships more precisely. However, it also leads to a great increase of computation consumption by hundreds of times, especially for high-resolution images. Thus, to illustrate the effectiveness of the proposed stitching framework, a conventional method based on the research done by Brown and Lowe (2007), developed by Microsoft Research Computational Photography Group, namely Image Composite Editor (ICE), is employed for the comparison. ICE is an advanced panoramic image stitcher. Given a set of overlapping photographs of a scene shot from a single camera location, the app creates high-resolution panoramas that seamlessly combine original images. ICE is widely used for various kinds of applications, especially by cameramen. The stitching result by ICE is shown in Fig. 12.

Because there is no quantitative index like accuracy in the evaluation step of image stitching (Lin *et al*. 2015), the comparison in this case study consists of two part: processing time and effect of key object stitching. To compare the processing more accurately, the algorithm proposed by Brown and Lowe (2007) is also written in Python, same as the proposed method. The resolution of images is 3684×5472. The implementation environment is based on an Intel Xeon CPU E5 (2.10 GHz×16). The processing time without image fusion is 0.368 seconds for the proposed method, and 10.952 seconds for the conventional approach. The processing time for image fusion in the proposed method is 5.947 seconds. Comparing to the conventional stitching process, this approach saves about 96.6% of computation consumption. Meanwhile, other supporting algorithms in this study need extra power with an increase of 54.3%. In a summary, there is about 42% reduction which means faster processing of the proposed method. ICE is developed based on Windows in C language and integrated with many other functions, and the processing time is longer than 11 seconds. The conventional method by Brown and Lowe (2007) has limitations for moving object stitching and would generate results like Fig. 11(b). Through Fig. 12, ICE may integrate this kind of function, because the left two moving vehicles are stitched well without ghosting, however the situation is not good for the right one. The reason for this may be that the velocity of the right vehicle is too big. While in Fig. 13, the proposed method can handle both situations well enough. Meanwhile, because the proposed method filters the feature points in the roadside region, the road in stitching result is straighter than that in ICE.

Considering that the panorama misses some objects because the UAV horizontal velocity generally increases and converges to 0 at the beginning and end of the flight, the scenes corresponding to the start and finish times are not selected. Based on the full-site panorama, vision-based object detection via deep-learning-based computer vision techniques can be performed to recognize ROIs (regions of interest) in the construction sites for the purpose of safety assessment and management, e.g., workers, safety helmet, etc., which is a follow-up study in the near future.

4. Conclusions

This study proposes a full-field high-resolution panorama generation method for civil engineering surveillance by camera-mounted UAV based on pre-registration using GPS information and efficient registration using feature points filtering, seeking out an efficient, prompt and accurate way for the combination of enormous visible images acquired by UAVs. It generates full-site panoramas and accomplishes the global surveillance using a rapid-processing ORB algorithm for feature points extraction and matching. Blurring and ghost effects caused by UAVs' continuous position changes in the flight plane are solved by optimal seaming line seeking, which ensures the accuracy and robustness of image stitching. The proposed approach saves more than about 40% cost comparing to the conventional process. The main contribution of this study is two-fold. The proposed novel image stitching method, on one hand, is based on pre-registration using GPS information, on the other hand, is based on specific feature point filtering method based on feature regions for registration, which reduces huge amounts of calculation cost and improves the processing speed. The proposed method can generate a full-site high-quality panorama in a much faster and more accurate manner, which is significant for the global safety control of construction management applications.

Acknowledgments

The research described in this paper was financially supported by National Key Rand D Program of China [2018YFB1600202, 2018YFC0705606], NSFC [Grant No. 51678204, 51638007, 51922034] and Guangxi Science Base and Talent Program [Grand No. 710281886032].

References

Akbar, M.A., Qidwai, U. and Jahanshahi, M.R. (2019), "An evaluation of image-based structural health monitoring using integrated unmanned aerial vehicle platform", *Struct. Control Health Monitor.*, **26**(1), e2276. https://doi.org/10.1002/stc.2276

Bang, S., Kim, H. and Kim, H. (2017), "UAV-based automatic generation of high-resolution panorama at a construction site with a focus on preprocessing for image stitching", *Automat. Constr.*, **84**, 70-80. https://doi.org/10.1016/j.autcon.2017.08.031

Brown, M. and Lowe, D. (2007), "Automatic panoramic image stitching using invariant features", *Int. J. Comput. Vision*, **74**(1), 59-73. https://doi.org/10.1007/s11263-006-0002-3

Calonder, M., Lepetit, V., Strecha, C. and Fua, P. (2010), "BRIEF: Binary robust independent elementary features", *Profeedings of European Conference on Computer Vision*, **6314**, 778-792.

Carlos, E.T. (2008), "Gauss Krüger projection for areas of wide longitudinal extent", *Int. J. Geographi. Info. Sci.*, **22**(6), 703-719. https://doi.org/10.1080/13658810701602286

De Melo, R.R.S., Costa, D.B., Lvares, J.S. and Irizarry, J. (2017), "Applicability of unmanned aerial system (UAS) for safety inspection on construction sites", *Safety Sci.*, **98**, 174-185. https://doi.org/10.1016/j.ssci.2017.06.008

Ellenberg, A., Branco, L., Krick, A., Bartoli, I. and Kontsos, A. (2014), "Use of unmanned aerial vehicle for quantitative infrastructure evaluation", *J. Infrastruct. Syst.*, **21**(3), 04014054. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000246

Hackl, J., Adey, B.T., Woźniak, M. and Schümperlin, O. (2017), "Use of unmanned aerial vehicle photogrammetry to obtain topographical information to improve bridge risk assessment", *J. Infrastruct. Syst.*, **24**(1), 04017041. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000393

Ham, Y., Han, K.K., Lin, J.J. and Golparvar-Fard, M. (2016), "Visual monitoring of civil infrastructure systems via camera-equipped unmanned aerial vehicles (UAV): A review of related works", *Visualiz. Eng.*, **4**(1), 1. https://doi.org/10.1186/s40327-015-0029-z

Irizarry, J. and Costa, D.B. (2016), "Exploratory study of potential applications of unmanned aerial systems for construction management tasks", *J. Manage. Eng.*, **32**(3), 05016001. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000422

Kim, S. and Kim, S. (2018), "Opportunities for construction site monitoring by adopting first personal view (FPV) of a drone", *Smart Struct. Syst., Int. J.*, **21**(2), 139-149. https://doi.org/10.12989/sss.2018.21.2.139

Li, H., Zhang, Q.-L., Yang, B., Lu, J. and Hu, J. (2015), "Development and application of construction monitoring system for shanghai tower", *Smart Struct. Syst., Int. J.*, **15**(4), 1019-1039. https://doi.org/10.12989/sss.2015.15.4.1019

Li, M., Chen, R., Zhang, W., Li, D., Liao, X., Wang, L., Pan, Y. and Zhang, P. (2017), "A stereo dual-channel dynamic programming algorithm for uav image stitching", *Sensors*, **17**(9), 2060. https://doi.org/10.3390/s17092060

Lin, C., Pankanti, S., Ramamurthy, K. and Aravkin, A. (2015), "Adaptive as-natural-as-possible image stitching", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, June.

Liu, P., Chen, A.Y., Huang, Y.-N., Han, J.-Y., Lai, J.-S., Kang, S.-C., Wu, T.-H., Wen, M.-C. and Tsai, M.-H. (2014), "A review of rotorcraft unmanned aerial vehicle (UAV) developments and applications in civil engineering", *Smart Struct. Syst., Int. J.*, **13**(6), 1065-1094. https://doi.org/10.12989/sss.2014.13.6.1065

Matthew, B. and Lowe, D.G. (2007), "Automatic panoramic image stitching using invariant features", *Int. J. Comput. Vision*, **74**(1), 59-73. https://doi.org/10.1007/s11263-006-0002-3

Metni, N. and Hamel, T. (2007), "A UAV for bridge inspection: Visual servoing control law with orientation limits", *Automat. Constr.*, **17**(1), 3-10. https://doi.org/10.1016/j.autcon.2006.12.010

Microsoft research image composite editor. http://research.microsoft.com/en-us/um/redmond/groups/ivm/ice/

Reagan, D., Sabato, A., Niezrecki, C., Yu, T. and Wilson, R. (2016), "An autonomous unmanned aerial vehicle sensing system for structural health monitoring of bridges", In: *Nondestructive Evaluation and Health Monitoring*, San Diego, CA, USA, March.

Rosten, E. and Drummond, T. (2006), "Machine learning for high-speed corner detection", *Profeedings of European Conference on Computer Vision*, **3951**, 430-443.

Rublee, E., Rabaud, V., Konolige, K. and Bradski, G.R. (2011), "ORB: An efficient alternative to sift or surf", *Proceedings of 2011 IEEE International Conference on Computer Vision*, Barcelona, Spain, November.

Torge, W. and Müller, J. (2012), *Geodesy*, Walter de Gruyter.

Xiang, R., Sun, M., Jiang, C., Liu, L., Zheng, H. and Li, X. (2014), "A method of fast mosaic for massive uav images", *Land Surface Remote Sensing II*, 9260, 92603W. https://doi.org/10.1117/12.2069201

Xu, X., Soga, K., Nawaz, S., Moss, N., Bowers, K. and Gajia, M. (2015), "Performance monitoring of timber structures in underground construction using wireless smartplank", *Smart Struct*. *Syst*., *Int*. *J*., **15**(3), 769-785. https://doi.org/10.12989/sss.2015.15.3.769

Zhang, W., Li, X., Yu, J., Kumar, M. and Mao, Y. (2018), "Remote sensing image mosaic technology based on surf algorithm in agriculture", *EURASIP J*. *Image Video Process*., **2018**(1), 85. https://doi.org/10.1186/s13640-018-0323-5

BS