

RESEARCH

Open Access



Remote sensing image mosaic technology based on SURF algorithm in agriculture

Weiping Zhang^{1,2}, Xiujuan Li², Junfeng Yu², Mohit Kumar³ and Yihua Mao^{3*}

Abstract

The remote sensing technology of unmanned aerial vehicle (UAV) is a low altitude remote sensing technology. The technology has been widely used in military, agricultural, medical, geographical mapping, and other fields by virtue of the advantages of fast acquisition, high resolution, low cost, and good security. But limited by the flying height of UAV and the focal length of the digital camera, the single image obtained by the UAV is difficult to form the overall cognition of the ground farmland area. In order to further expand the field of view, it is necessary to mosaic multiple single images acquired by UAV into a complete panoramic image of the farmland. In this paper, aiming at the problem of UAV low-altitude remote sensing image splicing, an image mosaic technique based on Speed Up Robust Feature (SURF) is introduced to achieve rapid image splicing. One hundred fifty ground farmland remote sensing images collected by UAV are used as experimental splicing objects, and the image splicing is completed by the global stitching strategy optimized by Levenberg-Marquardt (L-M). Experiments show that the strategy can effectively reduce the influence of cumulative errors and achieve automatic panoramic mosaic of the survey area.

Keywords: Low-altitude remote sensing image, SURF operator, Unmanned aerial vehicle (UAV), Panoramic splicing

1 Introduction

At present, remote sensing images have become an important means of information acquisition. Through high-resolution remote sensing images acquired by satellites, air planes, and radars, the information required by people can be obtained through image analysis. With the development of social economy and the need of national defense construction, the demand for high-resolution remote sensing images, and basic geographic information in all sectors of the society is more and more urgent, and their current requirements are becoming higher and higher. The remote sensing data acquired by satellites, air planes, and radar alone is difficult to fully satisfy the needs of image data acquisition and processing. Compared with traditional aerospace remote sensing technology, UAV low-altitude remote sensing technology, as a new low-altitude remote sensing technology, has many advantages such as high flexibility, easy operation, high resolution, and low investment. It has become one of the methods to obtain remote sensing data

[1]. UAV remote sensing uses remote sensing application technology, telemetry and remote control technology, remote sensing sensor technology, communication technology, GPS positioning technology, and unmanned aerial vehicle technology to realize intelligent and automated rapid acquisition of space remote sensing information such as earthquake disaster areas, natural environments, and land resources, at the same time, complete data processing, modeling, and application analysis [2]. As the UAV has the characteristics of light weight, high structural strength, low cost of use and procurement, convenient maintenance, and high compatibility in hardware, the application of UAV to low-altitude remote sensing data acquisition is becoming more and more extensive. At present, the application of small UAV mainly focuses on the following aspects: disaster assessment, disaster relief, flood control, environmental monitoring and management, urban management, resource management, and assessment of land dynamic changes, and so on [3].

At present, the application of remote sensing imagery to agricultural work has become a trend in the development of remote sensing technology and agricultural engineering. It uses remote sensing imagery as a source

* Correspondence: m18221325651@126.com

³College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, China

Full list of author information is available at the end of the article

of information and provides support for optimal agricultural production decisions through the processing of remote sensing imagery information [4]. Due to the rapid growth of crops and rapid changes in production demand, it is often necessary to obtain images for a specified period of time. Compared with manned aerial remote sensing and satellite remote sensing, UAV can carry out tasks quickly, with flexible flight time, convenient and fast operation, and can ensure the acquisition of dynamic data. The remote sensing data of high resolution resource satellites return to a long time and poor timeliness, so it is impossible to obtain the specified range of data in a short time. In addition, the UAV remote sensing is more suitable for the image acquisition in the small area. The UAV can be better applied to the agricultural field because of the advantages of obtaining high spatial and temporal spatial image and low cost. UAV remote sensing has potential application value in agriculture. The introduction of UAV into agriculture can greatly improve the level of agricultural modernization, and it has a long-term significance for promoting the development of modern agriculture [5].

However, the remote sensing images acquired by UAV have the characteristics of small amplitude and large quantity [6]. To get the global information of the whole region, we must mosaic and synthesize UAV remote sensing images. Image mosaic is a technique that combines several images with overlapping parts (the images may be obtained at different times, different viewing angles or by different sensors) into a large-scale seamless high-resolution image [7]. There are many methods of image mosaic, and there are differences in different algorithm steps, but the general process is the same.

Generally speaking, it mainly includes the following five steps [8]: (1) image preprocessing. It is mainly the basic operation of digital image denoising, histogram processing, and creating image matching templates. (2) Image matching. Image matching is the key to image mosaic technology. It refers to using a certain matching strategy to find the corresponding position of the template or feature point in the reference image that is to be stitched, and then determine the transformation relationship between the two images. (3) Establish a transformation model. According to the corresponding relationship between template or image features, calculate the value of each parameter in the mathematical model, thus establishing a mathematical transformation model of the two images. (4) Uniform coordinate transformation. According to the established mathematical transformation model, the image to be stitched is converted into the coordinate system of the reference image, and the unified coordinate transformation is completed. (5) Fusion reconstruction. The overlapping regions of the image to be stitched are fused to obtain a seamless panoramic image that is splice and reconstructed. According to the requirements of different stitching purposes, stitching accuracy and stitching speed, the low-altitude remote sensing image splicing of UAV can be divided into four types: quick mosaic with seam, panoramic image mosaic, uncontrolled orthophoto mosaic, controlled orthophoto mosaic [9]. The comparison of the four types of mosaic technology is shown in Table 1.

The panoramic image mosaic can be used to form a large panoramic image map from multiple images acquired by UAV. It can realize fast and seamless image stitching, and the object deformation is small, the relative

Table 1 Comparison of four UAV low-altitude remote sensing image mosaic techniques

Name	Advantages	Disadvantages	Scope of application
Quick mosaic with seam	Fast mosaic speed	There is a clear seam in the image	Used for rapid inspection of UAV aerial image, simple change of survey area, and remote control of UAV remote sensing images.
Panoramic image mosaic	Fast and seamless mosaic can be achieved in the airstrip, and the deformation of the object is small, the relative position relationship is more accurate, and the overall mosaic speed is fast.		Application to emergency and disaster monitoring
Uncontrolled orthophoto mosaic	The mosaic results have high precision, and no mosaic cracks and distortions are found after mosaic. The images also contain spatial coordinate information.	The accuracy of the elevation information of the image is not high	Change detection for urban and rural areas can also be used for orthophoto production which is not easy to enter the area.
Controlled orthophoto mosaic	Greatly reducing the number of field control points, its clustered processors can quickly achieve image matching, DEM automatic generation and other operations, with the highest accuracy.	The production cycle is relatively long	It is used for 3D rapid modeling, resource monitoring with high timeliness, and dynamic monitoring of large-scale engineering construction.

position relationship is more accurate, and the overall mosaic speed is fast. Through late-stage hyperspectral image analysis and other processing, it can be used to estimate the planting area of agricultural crops, growth monitoring, and pest and yield monitoring.

Stitching panoramic images usually uses the classic SIFT (scale invariant feature transform) algorithm for image matching. In order to solve the interference of translation, rotation and scale change between the images to be stitched, the SIFT algorithm realizes the matching of the features of two different images by extracting the invariant feature points of the image, and forms a high quality panoramic image by matching a number of same name points to the UAV remote sensing image [10]. However, this algorithm has a large amount of computation, and it is impossible to accurately identify and extract feature points from images with less blurred edges or feature points, and it is not possible to clearly identify edges and outlines. At the same time, the extracted points, especially the corners of objects in the image, cannot be registered, and the accuracy of the extraction is not high. In addition, low-altitude remote sensing images are greatly influenced by the air flow, which tends to cause large tilt angles and irregular overlaps, resulting in large image distortion. In addition, the instability of the photographic posture results in the continuous splicing of cumulative errors, and in severe cases, the distortion of the local edges. In view of these characteristics, this paper adopts the image mosaic based on SURF features, and uses the UAV remote sensing

image data to verify the algorithm. SURF algorithm is an improved algorithm based on SIFT algorithm. In order to achieve the purpose of acceleration, SURF algorithm uses Harr wavelet instead of Gauss filter to integrate the original image. By using Hessian matrix, the robustness of feature points is increased. So The SURF algorithm has greatly improved the speed and stability of feature points.

2 Method-image feature extraction and mosaic algorithm

2.1 SIFT algorithm

SIFT algorithm is a feature matching algorithm proposed by Lowe D in 2004 [11]. It is based on the technology of invariant features and proposes a point feature registration algorithm that keeps invariant on the translation and rotation of images. A brief process of extracting feature points from images using SIFT algorithm is shown in Fig. 1.

1. Detect scale space extremes. That is, the Gaussian convolution kernel is used to transform the original image to obtain the representation sequence in multi-scale space, where the calculation formula is [12]:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

Among them, (x, y) is the gray value of the image, $G(x, y, \sigma)$ is a scale-variable Gaussian function, which is defined as $G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$, and σ is a Gaussian scale factor.

Finally, the spatial extremum is extracted from these sequences for feature extraction.

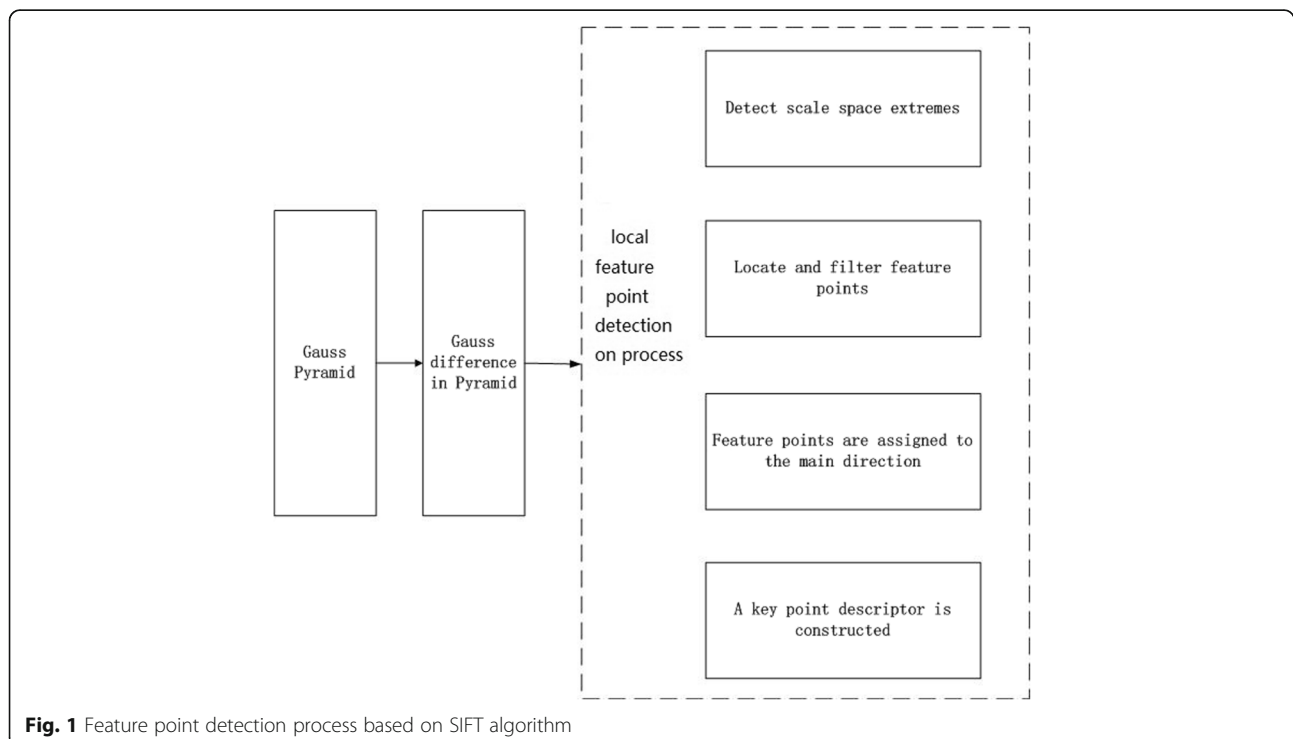


Fig. 1 Feature point detection process based on SIFT algorithm

2. Locate and filter feature points. Fitting the Laplacian scale space function $D(x, y, \sigma)$ in a two-dimensional continuous space. The Taylor formula is developed at the local extremum point (x_0, y_0, σ_0) , and the derivative of the expansion formula is equal to 0. The position coordinate \hat{x} of the extremum point is obtained, substituting the coordinate position of the extreme point into the scale space function obtains the extremum formula of the original function:

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial X} \hat{x} \quad (2)$$

When the offset in any direction is greater than 1/2, it indicates that the center point of the interpolation has been shifted to the neighboring point, then the extreme point of the class is discarded. Finally, in order to enhance the anti-noise ability and improve the stability, the Hessian matrix is used to eliminate unstable edge response points.

3. Feature points are assigned to the main direction. After determining the feature points, the main direction needs to be determined for each feature point so that it has rotation invariance. The calculation formula of the main direction of feature points is [13]

$$\begin{cases} m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \\ \theta(x, y) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \end{cases} \quad (3)$$

Where $m(x, y)$ denotes the gradient modulus value of the feature point (x, y) , $\theta(x, y)$ denotes the gradient direction of the feature point, and $L(x, y)$ denotes the Gaussian image of the scale on which the feature point is located. For each feature point, a histogram can be used to count the gradient distribution of the pixel gray values in the neighborhood of the pixel center, that is, determine the main direction of the feature point.

In addition, in order to make up for the instability of the feature points without affine invariance, the gradient amplitude of each sampling point should be weighted for each gradient histogram, and the weight of each sampling point is finally determined by the gradient modulus value of the sampling point and the Gaussian weight.

4. A key point descriptor is constructed. Through the calculation of the above three steps, each feature point detected contains three messages $((x, y), \sigma, \theta)$. That is, position, scale, and direction. Because the descriptor of the feature point is related to the scale of the feature point, the generation of the feature point descriptor needs to be carried out in the Gaussian pyramid space of the corresponding scale [14]. First, the neighborhood centered on the feature point is divided into $B_p \times B_p$ sub-blocks, and the edge size of each sub-block is $m\sigma$

pixels. The construction process of feature point descriptors is as follows: first, taking the feature point as the center, the rotation angle of the image in the neighborhood of the $(m\sigma(B_p + 1)\sqrt{2} \times m\sigma(B_p + 1)\sqrt{2})$ of the feature point is θ to the main direction of the feature point. Then, take the feature point as the center, select the $m\sigma B_p \times \sigma B_p$ size image block, and divide the interval into the $B_p \times B_p$ sub block, then use the gradient histogram to calculate the gradient accumulating value of all pixels in each sub block in eight directions, and form the seed point. A 128 dimensional feature vector is formed. In addition, in the process of constructing feature point descriptors, all the pixels in the neighborhood of the feature point need to be weighted by Gaussian, and all the pixels in the neighborhood range of the feature points need to be normalized two times in order to remove the influence of illumination and other factors [15].

2.2 Surf algorithm

The SURF feature based image square splicing is used in this paper. The basic process is image preprocessing, SURF feature matching, transformation parameter estimation, global splicing of L-M optimization, and image fusion.

2.2.1 Image preprocessing

Because of the uneven illumination of the original remote sensing image, the background gray level of the image is uneven. Therefore, in the binarization process of the image, in order to achieve the best effect, it is required to block the image. A threshold is determined by the valley on the gray histogram of the image, and then each block of the image is binarized.

2.2.2 SURF feature matching

The SURF algorithm uses the idea of SIFT algorithm. In order to improve the computing speed of feature extraction and matching, the SURF algorithm uses the approximate method of integral image and box filter, and keeps the image scale and rotation invariance as well as the better diversity and robustness. SURF feature matching is mainly divided into two steps: SURF feature extraction and SURF feature matching.

1) SURF feature extraction

- Construct scale space, extreme point detection. A convolution operation is performed on the integral image of the input image using an approximate Gaussian filter scaled up by layers to obtain a pyramid image. Calculate the Hessian matrix

determinant and get the feature point response value. By non-maximal suppression, each pixel in the scale space is compared with the other 26 pixels in the same layer and adjacent layers of the pixel to obtain local maxima and minima points.

- The Taylor expansion of the three-dimensional quadratic equation is used for surface fitting to achieve accurate positioning of interest points.
- The main direction of the feature point is determined. The main direction of the feature points in the SURF algorithm is determined based on Haar wavelet response and other information in the neighborhood of the feature points.
- Feature description vector generation. On the scale image of the feature point, the orientation of the coordinate axis and the orientation of the feature point are adjusted to be the same, and then the feature point is taken as the center. Similar to SIFT method of constructing the feature description vector, a 64-dimensional feature vector is obtained.

2) SURF feature matching

Here, the Euclidean distance of the SURF feature vectors of the two feature points is used as feature matching similarity measure. The characteristic points P_A and P_B feature vectors are D_{vA} and D_{vB} , respectively. The distance between two points is defined as [16]

$$D(P_A, P_B) = \sqrt{\sum_{i=1}^n (D_{vAi} - D_{vBi})^2} \quad (4)$$

The description of the process is as follows: the matching point search algorithm is used to find the feature points with the minimum and the second minimum distance to the match point. When the ratio of the minimum distance and the second minimum distance is smaller than the pre-set threshold, the matching is considered successful.

2.2.3 The transformation parameter estimation

The perspective transformation between images $X(x, y, 1)^T$ and $X'(x', y', 1)^T$ is expressed as

$$X' \sim HX = \begin{bmatrix} h0 & h1 & h2 \\ h3 & h4 & h5 \\ h6 & h7 & h8 \end{bmatrix} X \quad (5)$$

Since the extracted initial matching points generally have certain mismatched pairs, the RANSAC algorithm based on the perspective transformation constraints is required to first purify the matching pairs, and then the

least squares method is used to estimate the transformation matrix parameters. The process is to first calculate the transformation matrix H linearly by randomly selecting six points from the initial point pair, and then calculate the distance between other point pairs. A value less than the threshold is defined as an “inner point,” and a value greater than the threshold is defined as an “outside point.” Repeating iterations, selecting the set with the most “inner point” is the correct point pair after purification.

2.2.4 Global splicing of L-M optimization

In order to solve the cumulative error caused by the influence of photographic attitude, terrain fluctuation, and other factors [17], the global L-M optimized global splicing strategy is adopted in this paper. Based on the idea of minimizing the variance of the mean square distance between the same name points, this method uses the dynamic adjustment of each image transformation parameter in the splicing process to reduce the cumulative error effect and achieve the goal of global optimization.

The principle of this optimization strategy is to first select the reference plane, read the single images in turn, and then optimize the transformation matrix of each image to the reference plane at the same time so that the error of each image transformation to the reference plane is the smallest. For feature pairs (x_i, x_j) with the same name, x_i after the projection of the reference plane, the coordinate after projection to its adjacent image is x'_i and the distance difference is as follows:

$$d_{ij} = \|x_j - H_j^{-1} H_i x_i\| \quad (6)$$

Where H_i is the transformation matrix of the feature point of the image I_i projected onto the reference plane, H_j^{-1} is the transformation matrix from the reference plane back to the image I_j at this point.

According to the formula (6), summing up the distance difference of all other images that are overlapped with the current image I_i , we get the overall optimization target equation [18]:

$$e = \sum_{i=1}^n \sum_{j \in M(i)} \sum_{k \in F(i,j)} f(d_{ij}^k)^2 \quad (7)$$

Where n is the total image; $F(i,j)$ is a set of matching pairs between image I_i and image I_j ; d_{ij}^k is the distance difference calculated from the K matching point between image I_i and image I_j ; $M(i)$ is an image set that overlaps with I_i ; $f(x)$ is an error function, which is expressed as follows [19]:

$$f(x) = \begin{cases} |x||x| < x_{\max} \\ x_{\max}|x| \geq x_{\max} \end{cases} \quad (8)$$

The optimal transformation matrix is iteratively determined using the L-M algorithm [20], the basic process is described as follows:

1) SURF matching between all images, obtaining matching pairs, establishing an image adjacency relationship table.

2) Automatically select the reference plane. This paper uses the image I_i with the largest weight in the image sequence as the reference plane, the weight expression is [21]:

$$T = N + \frac{n}{S} \quad (9)$$

Where N is the number of overlapped images with the same name point pair as image I_i ; S is the area of the overlap area; n is the number of matching points for the overlapping area. The larger the T value, the larger the weight of the image I_i in the video sequence, which is taken as the reference plane.

3) The continuous transformation of the adjacent image transformation matrix in the airstrip is carried out to establish the general registration relationship between the photos and the reference image, which is used as the iterative initial value of the next L-M algorithm.

4) The best adjacent image is added in sequence according to the image adjacency relation table. In order to minimize the error when each image is transformed to the reference plane, the L-M algorithm is used to optimize the transformation matrix of each image to the reference plane.

5) Continue to add images, optimize transformation matrix, until all images are added, and finally output results.

2.2.5 Image fusion

When the geometric relationship between images is determined, in order to keep the visual color consistent and maintain a smooth visual transition, this paper uses a weighted average fusion method based on Gaussian model to stack multiple images into a panoramic image.

3 Experimental results and discussions

3.1 Experimental data

UAV remote sensing uses a low-altitude UAV as a platform. Different instruments are used to carry out telemetry on the surface according to different tasks to obtain the necessary surface information. Commonly used instruments include synthetic aperture radar, scanners, and CCD digital. In general, the UAV equipped with a CCD digital camera can acquire surface images flexibly, and the image accuracy is high. However, since low-altitude operations cannot obtain panoramic images in the target area, the captured images need to be stitched to form a panoramic image. This paper uses a Canon camera equipped with a 24 mm fixed focus camera on a UAV to shoot a rural farmland image data. The weather is clear and calm. The UAV is flying smoothly, and the image obtained is clear. Figure 2 shows a group of adjacent images from remote sensing images of rice fields obtained by UAV. It can be seen from the figure that there are more overlapping areas between two adjacent image pairs.



Fig. 2 One set of adjacent original images to be stitched



Fig. 3 Feature points extraction using SIFT algorithm and SURF algorithm

3.2 Comparison of image feature point extraction results

This paper compares and analyzes the feature point matching process and the splicing process of SIFT algorithm and SURF algorithm for adjacent images to be stitched. Figure 3 shows the effect of detecting feature points based on the SIFT algorithm and detecting feature points based on the SURF algorithm on adjacent images. The upper part of Fig. 3 is the image detection effect based on SIFT algorithm for detecting feature points. The lower part of Fig. 3 is the image detection effect based on SURF algorithm for detecting feature points. The results show that the SIFT extracts 580 feature points and takes 621.81 ms. The SURF extracts 356 feature points and takes 256.5 ms. From the feature extraction image, we can see that the points extracted by the SURF algorithm are significantly fewer than those extracted by the SIFT algorithm, and the quality of the extracted feature points is relatively high. It can be seen that the use of the SURF algorithm can eliminate invalid feature points outside the overlap region to some extent.

3.3 Mosaic image

In this paper, L-M algorithm is used to mosaic the panoramic images of the adjacent images as shown in Fig. 4. It can be seen from the figure that the image stitching is basically seamless, but the edges are partially missing. But the whole fitting effect is good. Because the data with large errors are eliminated, it is possible to eliminate the matching points with the same name. So the stitching effect is better.

4 Conclusions

This paper introduces the technology of remote sensing image mosaic based on SIFT algorithm and optimizes

the SIFT algorithm for the high resolution of the remote sensing image and too many detection feature points. The SURF operator is introduced into the low-altitude remote sensing image matching process, and the image mosaic technology based on the SURF algorithm is proposed. The difference between the detection feature points of SIFT algorithm and SURF algorithm is compared and analyzed by experiments. The SURF algorithm can eliminate the invalid feature points outside the overlapped region to a certain extent. It shows that it can improve the matching speed of the feature points greatly while maintaining the higher registration accuracy of the



Fig. 4 Panoramic view after mosaic

SIFT operator and has a strong performance in the low-altitude remote sensing image mosaic. In addition, a global mosaic strategy optimized by L-M is used to achieve panoramic mosaic of the survey area and better stitching results were obtained. Finally, the effectiveness of the SIFT algorithm and SURF algorithm for high resolution remote sensing image mosaic is compared.

Abbreviations

CCD: Charge-coupled device; GPS: Global Positioning System; L-M: Levenberg-Marquardt; SIFT: Scale invariant feature transform; SURF: Speed Up Robust Feature; UAV: Unmanned aerial vehicle

Acknowledgements

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

Funding

This work was supported by the National Natural Science Foundation of China (61662045) and the Special Program of talents Development for Excellent Youth Scholars in Tianjin.

Availability of data and materials

We can provide the data.

About the authors

Weiping Zhang was born in Shanghai, China, in 1982. He received the Ph.D. degree in computer science from Rostock University, Rostock, Germany. From 2014 to 2015, he was a Research Scientist with the Institute of BCRT in Berlin, Humboldt University, Berlin, Germany. He is currently a Associate Professor with the School of Information Engineering, Nanchang University, Nanchang, China. His research interests include machine learning, process information management systems, and real-time mobile measurements of physiological parameters.

Xiujuan Li was born in Shandong Province, China, in 1988. She received her master degree in Geographic Information System from Wuhan University. During her study period, her research interest focused on indoor and outdoor 3D visualization and location. Now, she works in Industrial Technology Research Institute of Zhejiang University, mainly studying wisdom water supply model and big data system.

Junfeng Yu was born in Tianjin, China, in 1987. He received the master's degree in communication and information system from Beijing Jiaotong University, Beijing, China. He has more than 6 years development and operation experience in the internet industry, particularly in data analysis and mining, domain name system, and routing technology. He is currently an R&D engineer in Binhai Industrial Technology Research Institute of Zhejiang University, Tianjin, China. His research interests include machine learning, data analysis, and data visualization.

Mohit Kumar is the Professor for "Computational Intelligence in Automation" in Faculty of Computer Science and Electrical Engineering, University of Rostock, Germany. He also holds the position of Research and Development Technical Director in Binhai Industrial Technology Research Institute of Zhejiang University, Tianjin, China. His research vision is to deliver a rigorous mathematical framework of intelligent computing algorithms with applications in machine learning, data analysis, signal modeling and analysis, image matching and mining, process identification, and optimization to fully utilize the potential of artificial intelligence paradigm for addressing challenging problems of real-world.

Yihua Mao, PhD, Professor. He was born in Zhejiang Province in 1963 and got a PhD degree in Zhejiang University. Now he serves in College of Civil Engineering and Architectural of Zhejiang University. He also serves as the deputy dean of Industrial Technology Research Institute of Zhejiang University and the dean of Binhai Industrial Technology Research Institute of Zhejiang University. His main research fields: the engineering economy and project cost, construction project management, system design, and large data analysis technology; enterprise strategy and technology innovation management.

Authors' contributions

All authors take part in the discussion of the work described in this paper. The author WZ, YM wrote the first version of the paper. The authors JY and YL did part experiments of the paper. MK revised the paper in different version of the paper. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Approved.

Consent for publication

Approved.

Competing interests

The authors declare that they have no competing interests. And all authors have seen the manuscript and approved to submit to your journal. We confirm that the content of the manuscript has not been published or submitted for publication elsewhere.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Electronic Information Engineering, Nanchang University, Nanchang, China. ²Binhai Industrial Technology Research Institute, Zhejiang University, Tianjin, China. ³College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, China.

Received: 3 July 2018 Accepted: 23 August 2018

Published online: 14 September 2018

References

- R. Hartley, P. Sturm, Computer vision and image understanding[J]. *Mach Learn Data Mining Methods Appl* **72**(2), 143–162(20) (1997)
- D.G. Lowe, Distinctive image features from scale-invariant keypoints[J]. *Int. J. Comput. Vis.* **60**(2), 91–110 (2004)
- F. Tao, M. Yokozawa, Z. Zhang, et al., Remote sensing of crop production in China by production efficiency models: Models comparisons, estimates and uncertainties[J]. *Ecol. Model.* **183**(4), 385–396 (2005)
- K. Shen, W. Li, Z. Pei, et al., Crop area estimation from UAV transect and MSR image data using spatial sampling method[J]. *Procedia Environ Sci* **26**, 95–100 (2015)
- L. Techy, D.G.S. Iii, C.A. Woolsey, Coordinated aerobiological sampling of a plant pathogen in the lower atmosphere using two autonomous unmanned aerial vehicles[J]. *J. Field Rob* **27**(3), 335–343 (2010)
- D.F. Schmidt, J. Botwinick, UAV-based imaging for multi-temporal, very high resolution crop surface models to monitor crop growth variability[J]. *Photogrammetrie - Fernerkundung - Geoinformation* **6**(6), 551–562 (2013)
- D.G.S. Iii, B.R. Dingus, C. Reinholtz, Development and application of an autonomous unmanned aerial vehicle for precise aerobiological sampling above agricultural fields[J]. *J. Field Rob* **25**(3), 133–147 (2008)
- H.P. Moravec, *Rover visual obstacle avoidance[C]// international joint conference on artificial intelligence* (Morgan Kaufmann Publishers Inc, 1981), pp. 785–790
- C.J. Harris, A combined corner and edge detector[J]. *Proc Alvey Vision Conf* **1988**(3), 147–151 (1988)
- D.G. Lowe, Distinctive image features from scale-invariant key-points. *Int. J. Comput. Vis.* **60**(2), 91–110 (2004)
- Y. Ke, R. Sukthankar, in *PCA-SIFT: a more distinctive representation for local image descriptors*. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004 (IEEE, 2004)
- H. Bay, A. Ess, T. Tuytelaars, et al., Speeded-up robust features[J]. *Comp Vision Image Underst* **110**(3), 404–417 (2008)
- T.W.R. Lo, J.P. Siebert, Local feature extraction and matching on range images: 2.5D SIFT[J]. *Comp Vision Image Underst* **113**(12), 1235–1250 (2009)
- W. Cheung, G. Hamarneh, N-SIFT: N-dimensional scale invariant feature transform for matching medical images. *IEEE international symposium on biomedical imaging: from Nano to Macro*. *IEEE* **18**(9), 720–723 (2007)

15. T. Suzuki, Y. Amano, T. Hashizume, Vision based localization of a small UAV for generating a large mosaic image[C]// Sice conference 2010, proceedings of. IEEE **9**(10), 2960–2964 (2010)
16. F. Fan, H. Wang, J. Li, et al., Novel averaging window filter for SIFT in infrared face recognition[J]. *Chin. Opt. Lett.* **9**(8), 081002–081002 (2011)
17. J. Luo, O.A. Gwun, Comparison of SIFT, PCA-SIFT and SURF[J]. *Int J Image Process* **3**(4), 143–152 (2009)
18. P.S. Addison, *The Illustrated Wavelet Transform Handbook*[J]. *Science Engineering Medicine & Finance* (Institute of Physics Publishing, 2016)
19. R. Szeliski, Video mosaics for virtual environments[J]. *IEEE Comp Graph Appl* **16**(2), 22–30 (1996)
20. Q.J. Yang, X.M. Xiao, Improved SIFT algorithm based on canny feature points[J]. *Comp Eng Des* **32**(7), 2428–2427 (2011)
21. T. Lindeberg, Scale-space for discrete signals[J]. *Pattern Anal Mach Intell IEEE Transact on* **12**(3), 234–254 (1990)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
