Automatic mosaic method of remote sensing images based on machine vision

S.P. Gao [1], M. Xia [1,2], S.J. Zhang [3]
[1] School of Mechanical and Electrical Engineering, Quanzhou University of Information Engineering, 362000, China, Quanzhou, Fengze District, Bodong Road, No. 249 ;
[2] Jiangsu Key Laboratory of Big Data Analysis Technology, Nanjing University of Information Science and Technology, 210044, China, Nanjing, Ningliu Road, No.219;
[3] College of Engineering and Technology, Yang-En University, 362014, China, Quanzhou, Luojiang District, Majia Town

Abstract

Unmanned Aerial Vehicle (UAV) remote sensing is a commonly used technical means in modern science and technology, but currently, remote sensing images captured by UAVs need to be spliced to obtain more comprehensive information. However, current image stitching techniques generally have shortcomings such as a small number of extracted features, low matching accuracy, and poor stability. To address the above issues, this study proposes an improved remote sensing image mosaic model on the bias of the Scale Invariant Feature Transform (SIFT) algorithm. Firstly, in this study, aiming at the problem that traditional SIFT cannot meet the requirements of feature extraction and matching for unconventional remote sensing images and special texture images, normalized cross correlation (NCC) and Forstner operator are introduced to optimize it, namely, a SIFT-NCC model is constructed. On this basis, for remote sensing images with high resolution and a wide range, this study further proposes a remote sensing image automatic mosaic model that combines point features and line features. That is, a linear segment detector (LSD) is introduced to extract the line features of remote sensing images. The experimental verification results of the final SIFT-NCC-LSD show that the matching accuracy for remote sensing images with different characteristics can reach over 95 %. Therefore, SIFT-NCC-LSD has good applicability.

Keywords: machine vision; remote sensing images; automatic splicing; point characteristics; line features; SIFT.

Citation: Gao SP, Xia M, Zhang SJ. Automatic mosaic method of remote sensing images based on machine vision. Computer Optics 2024; 48(5): 705-713. DOI: 10.18287/2412-6179-CO-1383.

Introduction

The vigorous development of contemporary science and technology has led to the continuous progress of high and new technologies has also arisen. It includes unmanned aerial vehicle remote sensing, satellite multi-angle remote sensing, and other earth observation technologies [1]. Remote sensing technology is a long-distance, non-contact target detection technology. Most remote sensing images contain rich geospatial and geological information, which have good application effects in many fields such as geological exploration and disaster detection. However, the commonly used unmanned aerial vehicle remote sensing technology currently obtains images with a large number of shortcomings, such as small range [2]. Therefore, for the detection of target regions, multiple remote sensing images need to be spliced to get more comprehensive information. However, in the process of data acquisition, the collected image data may be distorted due to factors such as terrain [3]. This situation can easily lead to problems such as layered remote sensing image stitching and low stitching accuracy. Therefore, the automatic mosaic technology of remote sensing images has become a key link in remote sensing data processing. Image stitching technology is also an important component of machine vision technology [4]. Machine vision mainly extracts information from objective images by simulating human visual functions, and then processes the extracted information for practical detection and control. Machine vision has the advantages of being fast, efficient, informative, and versatile [5]. The splicing process of remote sensing images is implemented based on machine vision. The current remote sensing image mosaic technology has shortcomings such as low number of features, low matching accuracy, and poor stability, which have a significant impact on the overall mosaic effect of remote sensing images. Therefore, based on the field of machine vision, this study proposed the SIFT-NCC-LSD model to improve the accuracy of remote sensing image stitching.

1. Related work

Machine vision is an interdisciplinary field that involves multiple fields. It mainly simulates human visual functions through computers, and then extracts relevant information from target images. This technology has a wide range of applications in image processing, target detection, equipment measurement and control. Recently, the artificial intelligence has made great progress in machine vision. Huang and Wu [6] used the Scale-invariant feature transform (SIFT) algorithm in machine vision to match images to achieve the capture of images from different angles. This method applied the image retrieval system of Hadoop to alleviate the shortcomings of MapReduce technology in ASIFT's high computing power. Experimental results verified that the proposed meth-

od is more effective than other advanced methods in processing multiple data sets. To improve the fusion ability of overexposed and underexposed industrial welds, Zheng et al [7] proposed a novel multi-view image fusion algorithm based on the Rotational Invariant Feature Transform (RIFT) algorithm. The method adopted autoencoder network structure, which had the characteristics of light structure and strong generalization ability. The proposed method could achieve the best parameters in terms of peak signal-to-noise ratio (PSNR) and root mean square error (RMSE) on the industrial weld data set. Deng and other researchers [8] proposed a visual synchronous positioning and mapping method combining Superpoint and K-Means algorithms, and applied it to the field of minimally invasive surgery. This method could extract the feature information of the image in lumen environment and obtain the parallax map by stereo matching method. Simulation experiments verified that the proposed method could increase the proportion of effective matching points and had good practicability. Ghorbani et al. [9] applied machine vision technology and image processing methods to the agricultural field to achieve accurate monitoring of planting, weeding, irrigation, and plant growth. Experiments showed that intelligent agricultural technology based on machine vision and image processing can reduce agricultural production costs and protect environmental resources. This study provided an outlook for the application of machine vision in the agricultural field, which is of practical significance. To enhance the intelligence of architectural decoration aesthetics, researcher Gong [10] constructed an aesthetic analysis system based on VR technology and machine vision, and applied it to the analysis of architectural decoration aesthetics. The system mainly applied a three-dimensional space detection algorithm to identify architectural decorative features and adjust the structure of the decoration in a timely manner. Finally, it was obviously that the established system could operate effectively and meet existing requirements.

The development of remote sensing image mosaic technology has been widely concerned by many researchers in recent years. Zhang and other scholars [11] proposed an enhanced algorithm for extracting feature points based on Hessian matrix in view of the shortcomings of traditional algorithms in existing remote sensing image mosaic techniques such as large computational complexity and low matching accuracy. This algorithm performed image preprocessing through image binarization, extracts feature points, and calculated normalized grayscale differences and two steps in the domain. After verification, the proposed algorithm had fast computing speed, good robustness and stability. Researchers such as Xu [12] introduced a remote sensing image mosaic method based on grid guided deformation and ground constraints to overcome the problem of inaccurate matching in traditional methods. This method used a new energy function to represent the deformation features of an image. The experiment was validated on a real background drone remote sensing image dataset. The results showed that the splicing system had a more natural splicing effect. Zhu et al. [13] found that traditional fast orientation and rotation algorithms had the disadvantage of high mismatch rate in the splicing process of remote sensing images. Therefore, an experimental remote sensing image mosaic method based on principal component invariant feature transformation was proposed. This method used a directional fast rotation method to obtain uniformly distributed feature points, and then used a random matching consistency algorithm to remove mismatched feature points. Eight sets of simulation experiments showed that this method had high image stitching efficiency and good stitching effect. To solve the problem of building artifacts in unmanned aerial vehicle remote sensing images, scholars such as Xu [14] constructed a remote sensing image mosaic method based on multi region guided local projection deformation. This method adapted to the local homography of image matching by gridding the image, and could calculate local coherence by weighting different regions of the image. Implementation validated the effectiveness of the proposed method through some cases. Zhang et al. [15] scientific research team found that current UAV aerial remote sensing image mosaic algorithms still lack real-time, robustness, and accuracy requirements. Aiming at the above problems, a novel real-time UAV image mosaic framework was experimentally constructed. This framework achieved real-time mosaic of unmanned aerial vehicle remote sensing images by introducing key frames during the tracking task. The constructed framework could obtain mosaic images in real-time and has high robustness.

In summary, both machine vision and remote sensing image stitching technology have achieved certain innovation and development in recent years. However, there is still a problem of poor splicing accuracy in practical applications. Therefore, this research is based on the field of machine vision, improving the traditional SIFT algorithm, and proposing a remote sensing image mosaic algorithm based on SIFT-NCC-LSD. The research aims to gather the advantages and characteristics of point features and line features, and improve the matching and stitching precision of remote sensing images.

2. Research on remote sensing image mosaic model based on improved SIFT extraction algorithm

2.1. Construction of remote sensing image mosaic model based on multiple feature points

In the field of remote sensing image mosaic, image feature points play a significant role. It can describe the key features of an image, that is, maximize the representation of the data information of the entire image in the form of minimum variables. The feature points of an image have good robustness against noise interference such as grayscale jumps in the image [16 – 17]. Scale invariant feature transform (SIFT) is a common feature extraction algorithm for remote sensing images, which is mainly used to detect and describe local features of images. SIFT

has good matching performance, mainly searching for extreme points in scale space, and extracting information such as the location and scale of the extreme points [18]. In addition, SIFT can also provide better processing results for abnormal conditions such as light differences, object motion, and occlusion in remote sensing images. The traditional SIFT algorithm is mainly divided into three steps: feature extraction, sign description, and feature matching. The specific content includes "searching for key points, accurately locating key points, determining the main direction of key points, generating feature point descriptors, and feature point matching" [19]. Assuming the existence of scale space $L(x,y,\sigma)$ and scale variable function $G(x,y,\sigma)$, the result of image subtraction in adjacent Gaussian scale spaces is shown in Equation (1):

$$D(x,y,\sigma)=\left(G(x,y,k\sigma)-G(x,y,\sigma)\right)\times I(x,y)= \\ =L(x,y,k\sigma)-L(x,y,\sigma). \tag{1}$$

In Equation (1), σ represents the scale space factor; k is a constant; $I(x,y)$ denotes a remote sensing image. Generate a Gaussian difference pyramid based on Equation (1), and then search for key points with larger extreme values in the Gaussian difference pyramid. The precise positioning equation for key points is shown in Equation (2):

$$D(X)=D+\left(\frac{\partial D}{\partial X}\right)^{T}X+\frac{1}{2}X^{T}\frac{\partial^{2}D}{\partial X^{2}}X. \tag{2}$$

In Equation (2), D denotes the output value of $D(x,y,\sigma)$ at the key point; $X=(x,y,\sigma)^{T}$ represents the offset of the key. The above equation represents the specific process of removing edge effects and edge points using the Hessian matrix, which can better improve the stability and noise resistance of key points. After that, the gradient amplitude m and argument angle θ of the Gaussian pyramid image (x,y) can be calculated.

$$m(x,y)=\sqrt{y-1\left(L(x+1,y)-L(x-1,y)\right)^{2}+\left(L(x,y+1)-L(x,y-1)\right)^{2}}. \tag{3}$$

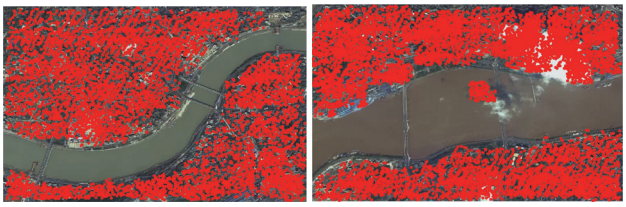
Equation (3) denotes the gradient amplitude of the remote sensing image (x,y), and its argument is shown in Equation (4):

$$\theta(x,y)=\arctan\frac{L(x+1,y)-L(x-1,y)}{L(x,y+1)-L(x,y-1)}. \tag{4}$$

After that, a feature point descriptor is generated based on the above equation and feature point matching is performed. The descriptor of SIFT is the overall result of gradient statistics on feature points and neighborhoods, and is a descriptor of local features. Fig. 1 shows the feature points extracted from a remote sensing image using the SIFT algorithm.

SIFT algorithm is a feature extraction and matching algorithm for remote sensing images based on scale space, which has strong applicability. However, in the real remote sensing image matching process, the SIFT algorithm cannot meet the requirements of feature extraction and matching for unconventional remote sensing images and special texture images, and there are often mismatches in the matching results. Therefore, this study proposes improvements to the traditional SIFT algorithm to enhance its feature matching performance. Normalized Cross Correlation (NCC) is a classic matching algorithm that mainly determines the degree of matching by calculating the crossing correlation values of template images and search images [20]. In view of the frequent mismatches in SIFT algorithms, this study applies the NCC algorithm to perform normalized cross correlation matching on the initial matching points shown in Fig. 1. That is, this study constructs an NCC-SIFT model. Equation (5) is the matching equation for the NCC algorithm:

$$\rho(r,c)=\frac{\sum_{x=1}^{p}\sum_{y=1}^{q}\left(g_{x,y}\bullet g'_{x+r,y+c}\right)-\frac{1}{p\bullet q}\left(\sum_{x=1}^{p}\sum_{y=1}^{q}g_{x,y}\right)\left(\sum_{x=1}^{m}\sum_{y=1}^{n}g'_{x+r}\right)}{\sqrt{\left[\sum_{x=1}^{p}\sum_{y=1}^{q}g_{x,y}^{2}-\frac{1}{p\bullet q}\left(\sum_{x=1}^{p}\sum_{y=1}^{q}g_{x,y}\right)^{2}\right]\left[\sum_{x=1}^{p}\sum_{y=1}^{q}g'^{2}_{x+r,y+c}-\frac{1}{p\bullet q}\left(\sum_{x=1}^{p}\sum_{y=1}^{q}g'_{x+r,y+c}\right)^{2}\right]}}. \tag{5}$$



a) *b)*

Fig. 1. SIFT algorithm extracting feature points. a) Left image, b) right image

In Equation (5), ρ represents the correlation coefficient of the NCC; g and g' represent pixel gray values in the left and right remote sensing images to be spliced, respectively; c,r represent displacement parameters per pixel, respectively. The SIFT algorithm optimized by the NCC algorithm has improved the mismatch situation to a certain extent, but it still causes error propagation due to texture and geometric distortions in the remote sensing image itself, thereby affecting the total mosaic effect of the image. Therefore, in this study, the Forstner operator is applied to eliminate errors in remote sensing images again. Forstner operator is a feature point extraction method with high efficiency and precision, which can find feature points with accurate parallax [21]. Taking the

left image in Fig. 1 as an example, the gradient of each pixel point is calculated, as shown in Equation (6):

$$\begin{cases} g_u = \dfrac{\partial g}{\partial u} = g_{i+1,j+1} - g_{i,j}, \\ g_v = \dfrac{\partial g}{\partial v} = g_{i,j+1} - g_{i+1,j}. \end{cases} \tag{6}$$

In Equation (6), g_u, g_v represent the gradient value of the pixel point. The covariance matrix equation of the Forstner operator is shown in Equation (7):

$$Q = N^{-1} = \begin{bmatrix} \sum g_u^2 & \sum g_{u g_v} \\ \sum g_{u g_v} & \sum g_v^2 \end{bmatrix}^{-1}. \tag{7}$$

In Equation (7), N represents the size of the side length of dividing the remote sensing image into several squares, that is, the area of each square is $N \times N$. After that, the interest value and pixel weight is calculated, as shown in Equation (8):

$$\begin{cases} q = \dfrac{4 Det\, N}{\left(tr\, N\right)^2} \\ \omega = \dfrac{1}{tr\, Q} = \dfrac{Det\, N}{tr\, N} \end{cases}. \tag{8}$$

In Equation (8), $Det\, N$ represents the determinant of the covariance matrix Q; $tr\, N$ indicates the trace of Q; q represents the value of interest; ω represents pixel weight. Thereafter, the threshold value is determined as shown in Equation (9):

$$\begin{cases} T_q = 0.5 \sim 0.75 \\ T_\omega = \begin{cases} f\overline{\omega}\left(f - 0.5 \sim 1.5\right) \\ c\omega_c\left(c = 5\right) \end{cases} \end{cases}. \tag{9}$$

In Equation (9), T_q and T_ω are given thresholds, respectively. If $q > T_q$, $\omega > T_\omega$, then the pixel is considered a candidate point. Finally, the Forstner feature points extracted from remote sensing images are combined with the feature points of the SIFT algorithm to generate a set of multiple feature points with the same name. Meanwhile, random sampling is used to further purify and eliminate mismatched points to obtain an accurate matching point set. The NCC matching process based on SIFT matching points and the NCC matching process based on Forstner feature points are shown in Fig. 2.

2.2. Construction of remote sensing image mosaic model based on point line complementary features

The method described in the previous section can achieve fully automatic intelligent splicing of remote sensing images, and has certain advantages in the amounts of matching points and splicing accuracy. It is suitable for splicing remote sensing images with many special textures and small visual angle changes [22]. For remote sensing images with high resolution and a wide range, only point feature stitching methods are prone to low matching accuracy or stitching failure [23]. Therefore, this section further proposes a remote sensing image automatic mosaic model combining point features and line features based on the above chapters. Firstly, based on the SIFT-NCC model proposed above, high-precision feature points are obtained as seed points for the method in this section. Subsequently, a homography matrix is calculated to determine the relationship between the spliced remote sensing images, as shown in Equation (10):

$$H = \begin{bmatrix} A & t \\ V^T & l \end{bmatrix} \quad A = \begin{bmatrix} a_1 & a_2 \\ a_4 & a_5 \end{bmatrix} \quad t = \begin{bmatrix} a_3 \\ a_6 \end{bmatrix} \quad V = \begin{bmatrix} a_7 \\ a_8 \end{bmatrix}. \tag{10}$$

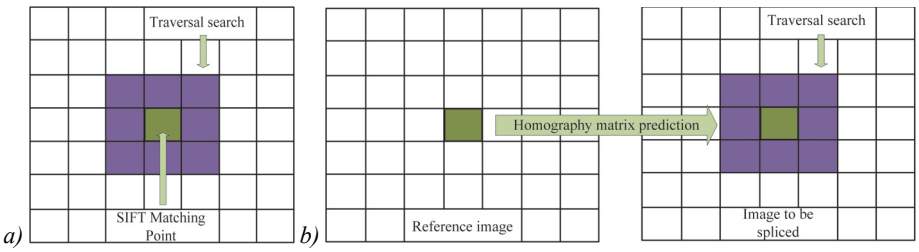


Fig. 2. Two matching processes. a) SIFT matching point NCC matching, b) forstner feature point NCC matching

In Equation (10), A, t, V represent the affine, translational, and perspective change parameters of remote sensing images, respectively. The root meaning square error shown in Equation (11) represents the stitching accuracy of remote sensing images:

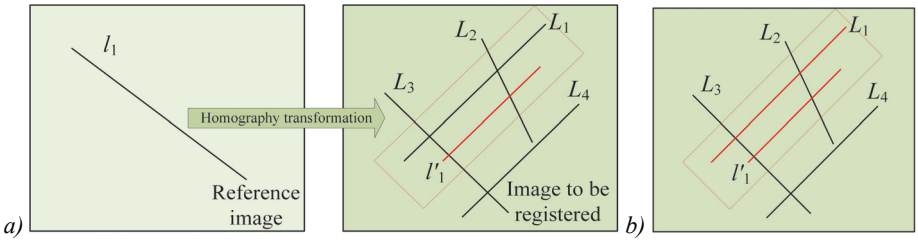
$$\varepsilon_{RMSE} = \sqrt{\sum_{i=1}^{n} \left\| x_i' - H x_i \right\|^2 / N}. \tag{11}$$

In Equation (11), N denotes the total number of feature points. The above operation achieves point feature matching of the image, followed by line feature matching. Linear Segment Detector (LSD) has a high accuracy and robustness for extracting linear features, so it is commonly used for linear feature extraction [24]. However, the linear features extracted by classical LSD methods are prone to clutter [25]. Therefore, to improve the accuracy of line feature matching, this study optimizes the LSD method. Firstly, a Gaussian template is used to denoise the original image and calculate the gradient amplitude and direction of the pixel. The gradient amplitude

is sorted to generate a linear support region, and a rectangular approximation is made to that support region. Then, the corresponding centerline of the region is taken as the initial straight line, and the final line features are obtained according to relevant principles. Coordinate the endpoints of each line feature, and eliminate or retain the line feature based on the set threshold value. To improve the matching accuracy, it is required to apply homomorphic geometric constraints and slope constraints to line features. Equation (12) is a specific expression for homomorphic geometric constraints:

$$\begin{cases} x' = \dfrac{a_1 x + a_2 y + a_3}{a_7 x + a_8 y + 1}, \\ y' = \dfrac{a_4 x + a_5 y + a_6}{a_7 x + a_8 y + 1}. \end{cases} \tag{12}$$

In Equation (12), (x,y) represents the coordinates in the left image to be spliced; (x',y') denotes the coordinates in the right image to be spliced. Fig. 3 shows the specific process of homomorphic geometric constraints and slope constraints. In Fig. 3, l_1 represents a straight line feature in the left image; L_1, L_2, L_3, L_4 represents line segment features in the right image, respectively. After performing homography prediction, l_1 determines a rectangle at the position of l_1' in the image on the right. The rectangle and L_1, L_2, L_3 both have intersections, so L_1, L_2, L_3 is the initial candidate line feature for the image. Meanwhile, L_4 is rejected because it does not intersect the rectangle. After geometric constraints are applied, the number of candidate line features obtained is still large, so slope constraints are applied to further filter the features. The slopes of the predicted line and the candidate line in the left and right images are calculated, and an appropriate threshold based on the slope of the predicted line is set. Compare the slope and threshold of candidate lines for culling or retention. In Fig. 3*b*, based on the slope calculated from the right image, the slope of L_1 of the candidate lines is retained due to the smallest difference between the slope of the matched line, and the rest are eliminated.



a) *b)*

Fig. 3. Constraint conditions. a) Homomorphic geometric constraint, b) slope constraint

After that, a descriptor is constructed for the selected feature lines. The obtained feature lines are constructed into a linear support region and divided into 5 equidistant bands. Construct the origin of the coordinate system from the midpoint of the line feature, and project the pixel gradient using Equation (13):

$$g' = \left(g^T \cdot d_\perp, g^T \cdot d_L\right)^T = \left(g'_{d\perp}, g'_{dL}\right)^T. \tag{13}$$

In Equation (13), g denotes the original gradient of the pixel; g' indicates the gradient in the coordinate system; $d_L\, d_\perp$ represents horizontal and vertical directions. After that, the influence of distance on gradient and edge effects need to be reduced. As shown in Equation (14):

$$\begin{cases} f_g(i) = \left(1/\sqrt{2\pi}\sigma_g\right) e^{-d_i^2/2\sigma_g^2} \\ f_i(k) = \left(1/\sqrt{2\pi}\sigma_i\right) e^{-d_k^2/2\sigma_i^2} \end{cases}. \tag{14}$$

In Equation (14), $f_g(i)$ represents the global weight; $f_i(k)$ represents local weights; σ represents the scale space factor; e is a constant. The resulting descriptor is shown in Equation (15):

$$LBD = \left(M_1^T, S_1^T, M_2^T, S_2^T, \ldots, M_m^T, S_m^T\right)^T. \tag{15}$$

In Equation (15), M_m^T represents the average vector of the stripe descriptor matrix; S_m^T represents the standard deviation vector of the stripe descriptor matrix. Therefore, Fig. 4 demonstrates the overall operation process of the remote sensing image automatic mosaic algorithm based on SIFT-NCC-LSD. First, point feature matching is performed, during which the coarse matching points obtained are refined using the NCC algorithm. After that, LSD algorithm is applied to extract complete linear features from the left and right images to be spliced, and homography constraints and slope constraints are used during the process. Finally, point and line features are fused to automatically splice remote sensing images.

3. Performance verification of remote sensing image mosaic model based on improved SIFT extraction algorithm

3.1. Performance verification of multi feature point remote sensing image mosaic model based on SIFT-NCC

To effectively verify the feasibility and effectiveness of the proposed algorithm, the experimental hardware platform selected in this study is a desktop computer with a main frequency of 3.3 GHz, 64 GB memory, 1080ti graphics card and i9 processor. The program is based on VisualStudio2010 C++ programming language and OpenCV and VLFeat computer vision library MFC algorithm program. To verify the effectiveness of SIFT-NCC based on multi point feature fusion, experiments first validate it. 1000 in 4 sets of pixels × A weakly textured aerial remote sensing image of 800 is taken as an experimental object, as shown in Fig. 5. There is rotation

change, that is, the image has a certain rotation angle; 5*b* There are scale changes, that is, images with translation, rotation, and scaling; Fig. 5*c* Lack of texture and obvious scaling; In Fig. 5*d*, there are obvious Angle changes, that is, images obtained from aerial photography from different angles. Based on this, the experiment set up three search domains to find the most suitable search domain for NCC matching. The three search domains are 1×1, 3×3 and 5×5. To better test the performance advantages of SIFT-NCC, experiments have been conducted to compare it with traditional SIFT algorithms and Speed up Robust Features (SURF) algorithms.

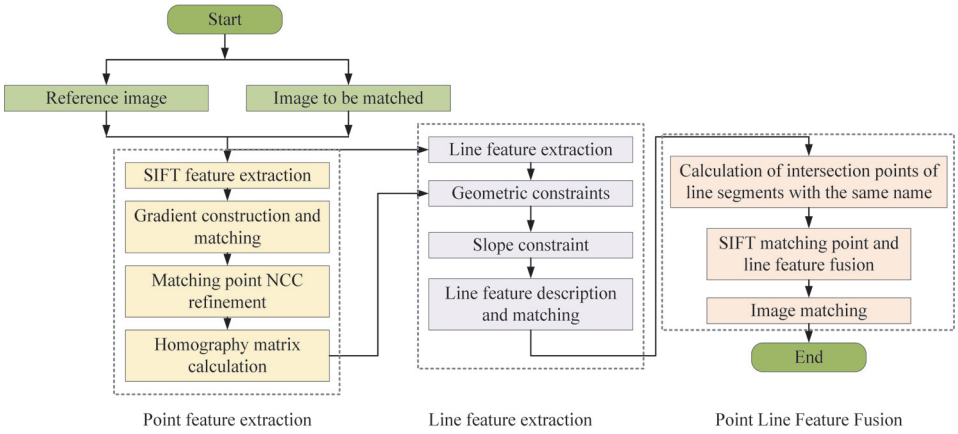


Fig. 4. Operation process of automatic remote sensing image mosaic algorithm based on SIFT-NCC-LSD

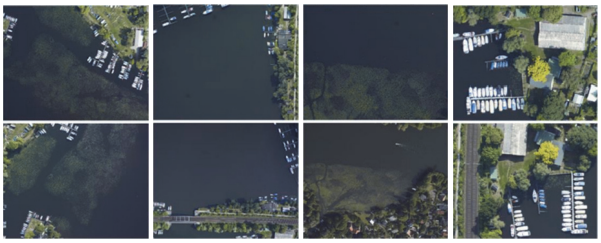


Fig. 5. Experimental original image data. a) Variation of spin, b) scale change, c) texture deficiency, perspective change

Fig. 6 shows the matching results of four sets of experimental images in three search domains. The highest matching accuracy is 3×3 pixel search domain, up to 98.66 %. Among the number of matching points in the three search domains, the total number of search fields obtained is the highest for a 5×5 pixel search field. In terms of the matching time spent: The search domain of 1×1 pixel takes the least time, 3×3 comes second; The search domain for 5×5 pixels takes the most time. Based on the above results, it is not difficult to find that the matching time is inversely proportional to the number of matching points obtained. The search domain of 3×3 pixels can obtain a larger number of matching points and a higher matching accuracy under a relatively suitable length of time. Therefore, the search domain of 3×3 pixels have a good matching effect and can meet the stitching accuracy required by this experiment.

Tab. 1 indicates the comparison of experimental outcomes of the three algorithms. From the accuracy of the algorithm, SIFT-NCC algorithm has great advantages. The accuracy rate of the SIFT-NCC algorithm for matching four groups of images is above 98 %, and its successful matching points for four groups of images are 1356 pairs, 723 pairs, 214 pairs, and 1012 pairs, respectively. In addition, compared to the other two algorithms, the SIFT-NCC algorithm has a lower mean square error and significantly superior multi feature point matching performance.

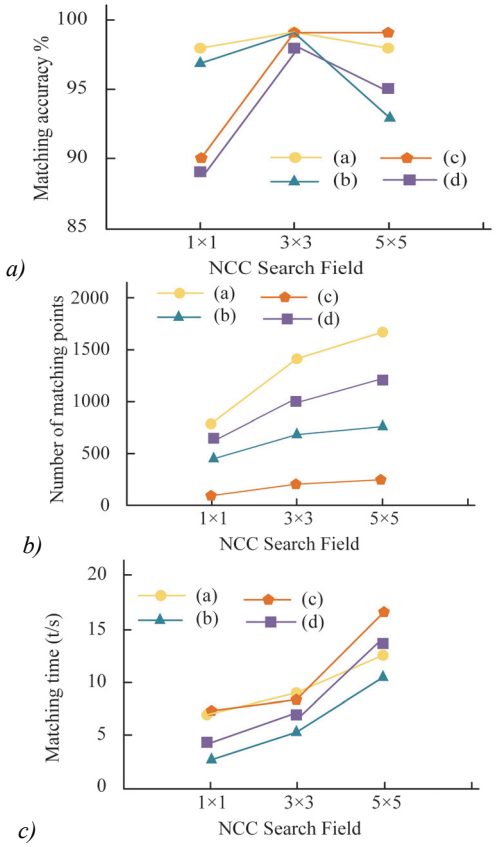


Fig. 6. Matching results of four groups of experimental images in three search domains. a) Matching accuracy of different search domains, b) number of matching points in different search domains, c) matching times for different search domains

Aggregate the data in Table 1 to get the results of the three algorithms shown in Fig. 7. In terms of the number of matching points, the SIFT-NCC algorithm can get the

maximum number of matching points for four groups of remote sensing images. For the accuracy of matching, the SIFT-NCC algorithm is still the highest among the three algorithms. Although the complexity of SURF algorithm is lower than that of SIFT algorithm, SURF is prone to more mismatches due to its low dimensionality of descriptors. In terms of matching time, the traditional SIFT algorithm takes the shortest time, followed by the SIFT-NCC algorithm, and the SURF algorithm takes the most time. The SIFT-NCC algorithm performs NCC and Forstner matching, so it takes more time. Combining all the results, SIFT-NCC algorithm has advantages in terms of the number of matching points, matching accuracy, and matching error, and its comprehensive stitching effect is the best.

Tab. 1. Comparison of experimental results of three algorithms

Image	SIFT		SURF		SIFT-NCC	
	N	Accuracy	N	Accuracy	N	Accuracy
5*a*	71	91.23 %	76	65.49 %	1356	99.36 %
5*b*	37	80.26 %	41	60.24 %	723	99.24 %
5*c*	6	/	2	/	214	99.12 %
5*d*	149	85.43 %	142	74.23 %	1012	98.45 %

3.2. Performance verification of remote sensing image automatic mosaic model based on SIFT-NCC-LSD algorithm

After constructing a remote sensing image mosaic model with multiple feature points, this study further proposes a remote sensing image mosaic model with complementary feature points and lines, namely the SIFT-NCC-LSD algorithm. Aiming at the characteristic of fusing point and line features, four groups of remote sensing images with line features are selected for the experiment. The four groups of images have rotation changes, scale changes, texture scarcity, and shading changes, respectively represented by A, B, C, and D. To verify the superiority of the performance of SIFT-NCC-LSD algorithm, experiments are conducted to compare it with traditional LSD algorithms.

Fig. 8 demonstrates the experimental results of line feature extraction. From the figure, the SIFT-NCC-LSD algorithm extracts fewer features than the traditional LSD algorithm. The reason is that the SIFT-NCC-LSD algorithm removes a line feature that does not match, thereby avoiding the interference of meaningless straight lines in the matching process. Therefore, the final line feature matching accuracy obtained by SIFT-NCC-LSD is high.

Tab. 2 shows the number of features extracted by SIFT-NCC-LSD algorithm and traditional LSD algorithm for four types of images and their error comparison. The features extracted by traditional LSD methods for four types of images are 1100 pairs, 241 pairs, 722 pairs, and 561 pairs, respectively; The features extracted by SIFT-NCC-LSD are 1036, 227, 699, and 533 pairs, respectively. However, the accuracy of SIFT-NCC-LSD is higher than that of LSD. Therefore, the usefulness of the constraints in the SIFT-NCC-LSD algorithm for removing invalid line features is verified.

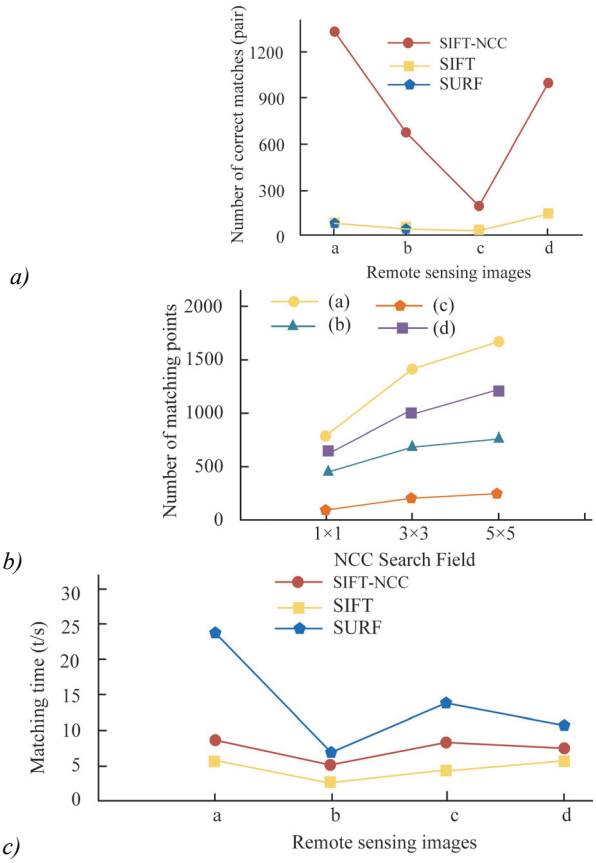
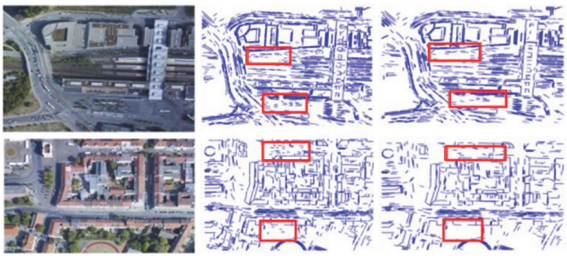


Fig. 7. Comparison of matching results of three algorithms. a) Comparison of the number of matching points. b) Matching accuracy comparison results. c) Matching time comparison results



a) b) c)

Fig. 8. Experimental results of line feature extraction using two algorithms. a) Original image, b) LSD, c) SIFT-NCC-LSD

Fig. 9 shows the matching results of the linear features of the SIFT-NCC-LSD algorithm, the SIFT-NCC algorithm, and the traditional SIFT algorithm. The feature matching degrees of the SIFT-NCC-LSD algorithm for four different remote sensing images are 648, 269, 374 and 501 pairs, respectively, which are higher than the other two algorithms. Observing the matching accuracy of the three algorithms for line features, it can be seen that SIFT-NCC-LSD can achieve a matching accuracy of more than 95 % for the four images, respectively 96.77 %, 98.74 %, 99.36 %, and 97.12 %. Although SIFT-NCC-LSD still suffers from incomplete online feature extraction, it cannot utilize constraints for line feature matching, im-

proving matching efficiency and accuracy. Traditional LSD and SIFT-LSD do not have relevant constraints, so they are prone to a large number of mismatches.

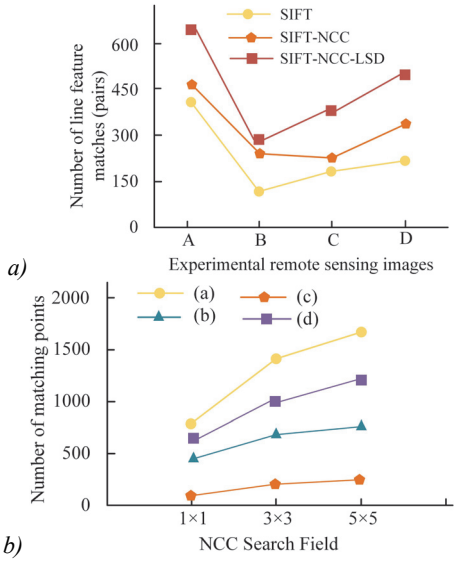


Fig. 9. Matching results of linear features of three algorithms. a) Comparison results of matching number of line features. b) Comparison results of line features matching accurasy

Fig. 10 shows the effect of point line feature fusion. SIFT-NCC-LSD has good applicability for remote sensing images with rotational changes, scale changes, texture scarcity, and shading changes. It can effectively extract point and line features when processing scenes such as blurred images, poor textures, and changes in brightness and shade.

Conclusion

During the mosaic process of remote sensing images captured by unmanned aerial vehicles (UAVs), it is easy to generate shortcomings such as a small number of extracted features, low matching accuracy, and poor stability. To address the above issues, this study first proposes a SIFT-NCC model. On this basis, a remote sensing image automatic mosaic model combining point and line features, namely the SIFT-NCC-LSD model, is further proposed. The experiment first verifies the performance of the SIFT-NCC model. It is clearly that the SIFT-NCC algorithm has a matching accuracy of over 98 % for four groups of remote sensing images with different features; The points successfully matched for four groups of images are 1356 pairs, 723 pairs, 214 pairs, and 1012 pairs, respectively; The quantity of matching points, matching accuracy, and matching time consumption of SIFT-NCC have advantages. Experimental results on the performance of SIFT-NCC-LSD show that the matching accuracy of SIFT-NCC-LSD for four different feature images can reach over 95 %, respectively 96.77 %, 98.74 %, 99.36 %, and 97.12 %. Therefore, the model constructed by the experiment has good applicability and effectiveness. There are still some shortcomings in this study. Future research can start from the uniform distribution of image feature points.

Tab. 2. Comparison of line feature extraction results between two algorithms

Image	LSD				SIFT-NCC-LSD			
	Total characteristics	Correct quantity	Accuracy	Error	Total characteristics	Correct quantity	Accuracy	Error
A	1100	1003	91.18 %	29.63	1036	1034	99.81 %	0.27
B	241	211	87.55 %	68.65	277	275	99.28 %	0.27
C	722	616	85.32 %	21.36	699	698	99.86 %	0.31
D	561	506	90.20 %	29.47	533	529	99.25 %	0.31

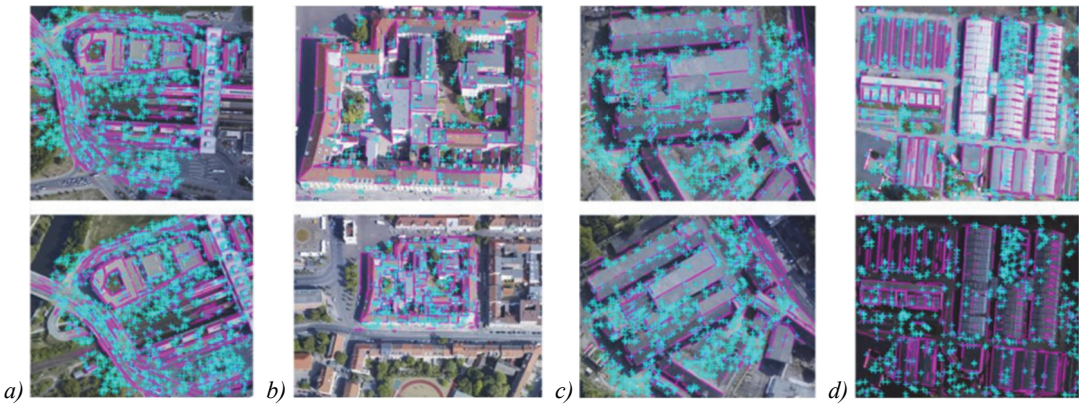


Fig. 10. Point line feature complementary fusion effect. a) Variation of spin. b) Scale change. c) Texture deficiency. d) Shading

Acknowledgements

The research is supported by: Educational Teaching Reform of Undergraduate Universities in Fujian Province-Reform of Practical Teaching Mode for Cultivating Local Applied Undergraduate Engineering Talents under the Background of Integration of Industry and Education (No.: FBJG20200355).

References

[1] Wu S, Wang Z, Shen B, et al. Human-computer interaction based on machine vision of a smart assembly workbench. Assem Autom 2020; 40(3): 475-482.

[2] Grigorev I, Shadrin A, Katkov S, et al. Improving the quality of sorting wood chips by scanning and machine vision technology. J For Sci 2021; 67(5): 212-218.

[3] Chiu T L, Lin S Z, Ahmed T, et al. Pilot study of a new type of machine vision-Assisted stereotactic neurosurgery for EVD placement. Acta Neurochir 2022; 164(9): 2385-2393.

[4] Paul S, Pati UC. A comprehensive review on remote sensing image registration. Int J Remote Sens 2021; 42(14): 5396-5432.

[5] Zhang J, Xu S, Zhao Y, et al. Aerial orthoimage generation for UAV remote sensing. Inf Fusion 2023; 89: 91-120.

[6] Huang Y, Wu H. Image retrieval based on ASIFT features in a Hadoop clustered system. IET Image Process 2020; 14(1): 138-146.

[7] Zheng Q, Zhao Y, Zhang X, Zhu P, Ma W. A multi-view image fusion algorithm for industrial weld. IET Image Process 2023; 17(1): 193-203.

[8] Deng L, Liu Z, Zhang T, Yan Z. Study of visual SLAM methods in minimally invasive surgery. Math Biosci Eng 2023; 20(3): 4388-4402.

[9] Ghorbani M, Aboonajmi M, Asefpour VK. The machine vision technology in precision agriculture: A comprehensive review on principles and applications. Soft Comput J 2021; 9(1): 92-113.

[10] Gong MC. Analysis of architectural decoration esthetics based on VR technology and machine vision. Soft Comput 2021; 25(18): 12477-12489.

[11] Zhang T, Zhao R, Chen Z. Application of migration image registration algorithm based on improved SURF in remote sensing image mosaic. IEEE Access 2020; 8: 163637-163645.

[12] Xu Q, Chen J, Luo L, et al. UAV image stitching based on mesh-guided deformation and ground constraint. IEEE J Sel Top Appl Earth Obs Remote Sens 2021; 14: 4465-4475.

[13] Zhu JT, Gong CF, Zhao MX, et al. Image mosaic algorithm based on PCA-ORB feature matching. Int Arch Photogramm Remote Sens Spat Inf Sci 2020; 42: 83-89.

[14] Xu Q, Chen J, Luo L, et al. UAV image mosaicking based on multi-region guided local projection deformation. IEEE J Sel Top Appl Earth Obs Remote Sens 2020; 13: 3844-3855.

[15] Zhang F, Yang T, Liu L, et al. Image-only real-time incremental UAV image mosaic for multi-strip flight. IEEE Trans Multimed 2020; 23: 1410-1425.

[16] Lauguico SC, Concepcion RS, Alejandrino JD, et al. A comparative analysis of machine learning algorithms modeled from machine vision-based lettuce growth stage classification in smart aquaponics. Int J Environ Sci Dev 2020; 11(9): 442-449.

[17] Soltani FM, Sardari H. Defect detection in fruit and vegetables by using machine vision systems and image processing. Food Eng Rev 2022; 14(3): 353-379.

[18] Kang Z, Yang J, Guo H. Automatic garbage classification system based on machine vision. J Zhejiang Univ (Eng Sci) 2020; 54(7): 1272-1280.

[19] Wang N, Zhang G, Ren L, et al. Novel monitoring method for belt wear state based on machine vision and image processing under grinding parameter variation. Int J Adv Manuf Technol 2022; 122(1): 87-101.

[20] Habib MT, Majumder A, Jakaria AZM, et al. Machine vision based papaya disease recognition. J King Saud Univ – Comput Inf Sci 2020; 32(3): 300-309.

[21] Pei T, Xu J, Liu Y, et al. GI Science and remote sensing in natural resource and environmental research: Status quo and future perspectives. Geogr Sustain 2021; 2(3): 207-215.

[22] Zeng Y, Ning Z, Liu P, et al. A mosaic method for multitemporal data registration by using convolutional neural networks for forestry remote sensing applications. Computing 2020; 102: 795-811.

[23] Chen B, Xia M, Qian M, et al. MANet: A multi-level aggregation network for semantic segmentation of high-resolution remote sensing images. Int J Remote Sens 2022; 43(15-16): 5874-5894.

[24] Song B. Optimization of progressive image mosaic algorithm in fine art image fusion for virtual reality. IEEE Access 2020; 9: 69559-69572.

[25] Lca B, Yza B, Cwc C, et al. Histograms of oriented mosaic gradients for snapshot spectral image description. ISPRS J Photogramm Remote Sens 2022; 183: 79-93.

Authors' information

Shanping Gao (b. 1985) master, studied at Huaqiao University. Associate professor in Quanzhou University of Information Engineering. His interest is mechanical engineering, digital design and manufacturing.
E-mail: *gsp@qziedu.cn*

Min Xia, (b. 1983), IEEE Member, received the Ph.D. degree in Cybernetics Control Engineering from Donghua University, Shanghai, China, in 2009. Currently, he is a Professor with the Nanjing University of Information Science and Technology, and the Deputy Director of Jiangsu Key Laboratory of Big Data Analysis Technology in Nanjing, China. His research interests include machine learning theory and its application, and graph structure data analysis.
E-mail: *xiamin@nuist.edu.cn*

Sujia Zhang (b. 1987) graduated from Huaqiao University in 2018, majoring in Electronic and Communication Engineering. Currently she works as the lecturer at the College of Engineering and Technology in YangEn University. Research interests are quantum reversible logic synthesis. Serve as the corresponding author of this article.
E-mail: *zhang_sj@yeu.edu.cn*

Received June 25, 2023. The final version – January 11, 2024.