Transmission line image mosaic based on improved SIFT-LATCH

Shilin Dong¹, Li Lu *¹
¹ Electrical Engineering, Shanghai DianJi University, Shanghai, 201306, China
*Corresponding author’s e-mail: lul@sdju.edu.cn

Abstract: In the transmission line inspection, in order to obtain wide-view and large-scale images, it is necessary to stitch images taken by drones. This article proposes an improved SIFT-LATCH-based UAV aerial transmission line image stitching algorithm. First, the image is processed in parallel through multi-threading. The SIFT algorithm is used to extract the feature points in the image, and the LATCH descriptor is used to describe the feature points. The sampling consensus algorithm is used to purify the feature points, and finally the translational homography matrix is used to solve the problem of black lines in the stitched image. Experiments show that the algorithm proposed in the article has the advantages of good real-time and high accuracy in splicing UAV aerial pictures of transmission lines.

1. Introduction
With the development of society and economy, electric power carries indispensable factors in all aspects, and the laying of a large number of electric equipment has increased the problem of aggravating line inspection work. At present, the inspection of electrical equipment based on aerial image processing has shown great potential in real-time and intelligence. At the same time, the use of drones for aerial inspection of transmission lines has become an important research direction [1]. In order to reduce the burden of inspectors and improve work efficiency, the images acquired by drones are used to stitch together a large-scale and wide-field seamless aerial transmission line image, so that it provides convenience for inspectors to discover and deal with hidden dangers in power transmission channels in time based on images. Due to the limited viewing angle of the drone's airborne camera and the influence of factors such as wind, it is difficult to obtain the transmission line image at the same height and the same viewing angle, which brings certain difficulties to the subsequent image stitching. Therefore, a splicing algorithm with good real-time and robustness is needed.

2. Image feature extraction based on SIFT

2.1 The construction of scale space.
Using Gaussian kernel function and convolution of the image to be spliced to build a spatial pyramid of aerial images:

\[ L(x, y, \sigma) = I(x, y) * G(x, y, \sigma) \]  (1)

In the formula, \( L(x, y, \sigma) \) represents the scale space image generated after Gaussian convolution operation, \( I(x, y) \) is the original image, and \( G(x, y, \sigma) \) is the Gaussian kernel function, as in formula (2) Shown:
In the formula, \( \sigma \) is the scale factor. The larger the value of \( \sigma \), the lower the pixel of the image, and the smaller the value of \( \sigma \), the higher the pixel of the image. The Gaussian difference scale space is formed by convolution of Gaussian difference kernels of different scales with the image:

\[
D(x, y, \sigma) = I(x, y) * (G(x, y, k\sigma) - G(x, y, \sigma))
\]

By establishing a multi-scale space, processing information at different scales can be obtained. If you want to include feature points at any scale and maintain the scale space unchanged, you need to build pyramids at different scales of the image. The adjacent towers are formed by sampling the space of the previous scale, and their size is a quarter of the image of the previous tower. The image description of the scale space is shown in Figure 1:

2.2. Feature point (extreme point) detection

After the Gaussian pyramid is constructed, the adjacent layers in the tower are calculated by the Gaussian difference method to form a pyramid difference scale image. Finally, the points after the scale space difference are compared with all the points in the neighborhood. If there is a certain point in the image if a point is a maximum or minimum point in this layer and all the neighborhoods corresponding to the next layer and the upper layer, the extreme point of this image is the point of interest in the scale space.

3. Improved SIFT-LATCH algorithm

3.1. Triple Binary Descriptor Establishment

The LATCH descriptor forms a bit string by calculating the comparison value of the pixel block in the window. The value of each bit in the bit string is determined by comparing the gray value of the pixel point pair in the relevant area of the feature point. The LATCH descriptor is a variant binary feature descriptor, and only needs to calculate the Hamming distance in the processing of image matching, thus saving the time of image matching.[2]

After comparing the F-norm of the pixel block in the neighborhood window of the feature point, part of the noise influence in the matching process is eliminated. With the feature point to be detected as the center of the circle, an image block of a specific size is selected, and \( T \) groups of image blocks are defined in each detection window \( W \). The specific formulas are shown in formula (4) and formula (5):

\[
S = \{S_t\}_{t=1,2,\ldots,T} = \{[P_{t,\sigma}, P_{t,1}, P_{t,2}]\}_{t=1,\ldots,T}
\]

\( P_{(t,\alpha)} \) is the pixel block, \( P_{(t,1)}, P_{(t,2)} \) are the associated pixel block. Calculate the F-norm of the matrix formed by the gray value difference of the corresponding pixels between \( P_{(t,1)} \) and \( P_{(t,\alpha)} \),
and between \( P_{(t,2)} \) and \( P_{(t,\alpha)} \). After the F-norm value is compared, the value of the corresponding bit in the binary string can be obtained [3].

\[
g(W, S_t) = \begin{cases} 1, & \|P_{t,\alpha} - P_{t,1}\|_F > \|P_{t,\alpha} - P_{t,2}\|_F \\ 0, & \end{cases}
\] (5)

The descriptor generated by the F-norm obtained by the comparison calculation reduces the influence of the noise in the image on the matching process, so that the algorithm has better robustness.

During program execution, the time advantage of binary descriptors is reserved by LATCH descriptors, which are an order of magnitude faster than descriptors based on histograms. In terms of stability, the LATCH descriptor further narrows the gap between the histogram-based descriptors. Compared with other binary descriptors, the LATCH descriptor has more advantages in most image sets.

3.2. Homography matrix

In general, the homography transformation can be understood as the mapping relationship between the object position in the world coordinate system and the pixel coordinate system. The matrix after the corresponding transformation is called the homography matrix. The homography matrix is defined as shown in formula (6): In the formula, \( u \) and \( v \) represent the coordinates in the pixel coordinate system; \( s \) represents the scale factor; \( f_x, f_y, u_0, v_0, \) and \( \gamma \) represent 5 camera internal parameters, and \( R \) and \( t \) represent external parameters. [4]

\[
H = s \begin{bmatrix} f_x & \gamma & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} = sM[r_1 \ r_2 \ t]  
\] (6)

Among them, \( M \) is the internal parameter matrix:

\[
M = \begin{bmatrix} f_x & \gamma & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}  
\] (7)

When the homography matrix is calculated, when stitching, the left image will move to the left when stitching, which will cause part of the image to disappear, causing the black line problem. Therefore, it is necessary to shift the image after the homography transformation to the right. Therefore, first, the left part of the spliced image is shifted to the right by an image width during the homography transformation, so that the corresponding position of the optimal pairing point on the original image and the matrix-transformed image can be obtained, thereby positioning the splicing point of the image.

4. Algorithm flow design

The main flow of the article algorithm is as follows:

1. Use multi-threaded parallel technology to process two images to be processed at the same time.
2. Use the improved SIFT-LATCH algorithm proposed in the article for feature extraction, after calculating and describing the feature points, pair the already described feature points.
3. The article adopts KNN algorithm for rough matching of feature points.
4. The random sampling consensus (RANSAC) algorithm is used to obtain the best transformation model, and the homography matrix parameters are calculated.
5. A weighted fusion algorithm is adopted to fuse the stitching gaps, so that the pixels in the overlapping area are superimposed according to the corresponding weights to eliminate the stitching gaps of the image and make the image transition more natural.

5. Experimental results and analysis

To measure whether the algorithm proposed in the article can be applied to image splicing of transmission lines, it is necessary to analyze the algorithm's feature point extraction speed, matching
speed and registration accuracy. Through experiments, the improved SIFT-LATCH image stitching algorithm and ORB algorithm, AKAZE algorithm, and SIFT algorithm were compared and analyzed in four aspects: interest point extraction, feature matching, matching accuracy and total time. The hardware of the experimental running environment is AMD Ryzen 5 3.85GHz, the memory is 8GB, the 64-bit WIN10 operating system desktop computer. All the algorithms in the experiment are implemented based on opencv3.4.1, the programming language is C++, and the programming environment is Visual Studio 2015.

The images used in the data set established in the article are all four image sets established by drone shooting, and each image set contains 20 images with a resolution of 3000×4000.

Figure 4 is an example of an experiment, where (a) represents an image with parallax, (b) represents a rotated image, and (c) represents an image with brightness changes.

![Experimental Images](image)

**Figure 2. Examples of experimental images**

5.1. **Average time spent in matching phase**
Experiments show that compared with the other three algorithms, the algorithm used in the article has a great improvement in the average time of the matching phase. Table 1 shows the time used for pairing of feature points after calculation and description.

| Image group | SIFT (t/s) | ORB (t/s) | AKAZE (t/s) | Article algorithm (t/s) |
|-------------|------------|-----------|-------------|------------------------|
| First group | 0.470      | 0.006     | 0.127       | 0.087                  |
| Second Group | 0.477      | 0.008     | 0.113       | 0.086                  |
| Third group | 0.513      | 0.008     | 0.121       | 0.087                  |
| Fourth group | 0.452      | ---       | 0.105       | 0.081                  |

It can be seen from Table 1 that in the feature point matching (feature pairing) stage, the improved SIFT-LATCH algorithm proposed in the article takes significantly less time than the SIFT algorithm. Among them, the ORB algorithm is the least time-consuming, but the algorithm is the worst in terms of stability.
After the program reads the image, image matching is mainly divided into three stages: (1) feature extraction; (2) feature calculation and description; (3) feature matching. Because there is no public data set of transmission line images, the experiment is carried out in the data set constructed in the article. The SIFT-LATCH algorithm, the SIFT algorithm, the ORB algorithm and the AKAZE algorithm are compared and analyzed. When the matching point is limited to 1500, the average time spent in each stage is compared, and the result is shown in Figure 3:

It can be seen from Figure 3 that in terms of time efficiency, the ORB algorithm is the fastest, followed by the algorithm proposed in the article, and the SIFT algorithm takes the longest time.

5.2. Matching accuracy comparison

Because each algorithm extracts and calculates and describes the feature points in different ways, the number of matching point pairs is very different when the feature points are not limited. The article uses the correct rate of matching as the criterion for evaluating the robustness of the algorithm. For the illumination difference and visual difference images, Fig. 4 show the histogram of the matching accuracy of the algorithm in the text compared with the SIFT, ORB, and AKAZE algorithms.

![Figure 3. Matching stage comparison result](image3.png)
![Figure 4. Illumination difference image contrast histogram](image4.png)

According to the comparison result in Fig. 4, when dealing with images with different illumination, the SIFT algorithm has the highest correct rate, followed by improved SIFT-LATCH and AKAZE algorithms. The ORB algorithm has poor stability.

Figure 5 contains (a), (b), (c) sub-pictures corresponding to the splicing results of the three groups of experimental pictures in Figure 2 (a), (b), (c).

![Figure 5. Algorithm splicing results in this paper](image5.png)

6. Conclusion

In the transmission line inspection, due to the limited viewing angle of the UAV, the wide-view and large-scale panoramic image cannot be obtained. In response to this problem, the article proposes an improved SIFT-LATCH algorithm based on multi-threaded acceleration to splice aerial transmission line images, the experimental results show that after the homography matrix is calculated, the image is moved by one distance, and after the weighted algorithm fusion processing, the problem of black lines generated in image mosaic is effectively removed.
In addition, the algorithm proposed in the article deals with when the image is illuminated. When the angle changes, it can still perform image stitching better. Compared with the traditional SIFT algorithm, it has more advantages in real-time performance and is of great significance in the intelligent inspection of aerial transmission lines.

Acknowledgments
Thanks to the tutor who gave me advice and guidance in the research of this article, and the colleague who gave me encouragement and help.

References
[1] P, Hong., Li, H.L., Du,Y., Ning,B., Guo,L.(2020) Stitching algorithm of aerial images taken by power patrol uav.J.Computer Applications and Software, 37(05): 243-248+255.
[2] LI, Z., LIU,J.Y., LI,H., Zhou, X.G., L,W.P.(2017)Feature detection and description algorithm based on ORB-LATCH.Journal of Computer Applications, 37(06): 1759-1762+1781.
[3] Song, W., Wang, Y.B.(2018) UAV image registration based on improved AKAZE algorithm J. Journal of Electronic Measurement and Instrumentation, 32: 96-102.
[4] Chi.,L.Y., Zhang,H., Zhao C.X.(2020) Research on image mosaic and positioning algorithm based on local homography matrix.J.Navigation Positioning and Timing, 7: 62-69.
[5] Huang,S., Wu,Z.S., Ren,G., Liu,H.J.,G,Y.(2020) Review of electric power intelligent inspection robot.J. ElectricalMeasurement & Instrumentation, 57: 26-38.