Research on image matching algorithm based on improved SIFT UAV

Rong Guo¹, Shixin Li², Ruyi Cai³ and Xiali Sun⁴
Tianjin University of Technology and Education, Tianjin, China

Corresponding author and e-mail: Shixin Li, li_shixin@sina.com

Abstract. Aiming at the low accuracy of UAV image registration algorithm, and the complex calculation of SIFT algorithm feature descriptor and the low time efficiency, an image matching algorithm based on improved SIFT is proposed. Firstly, the image of UAV is preprocessed, and the feature point operator is detected by Shi-Tomasi corner, then is extracted by SIFT feature vector, and the dimension of feature vector is reduced by PCA. Then the nearest neighbor method is used for feature matching, and the BBF algorithm is used to search the feature values to improve the matching speed. Finally, the RANSAC algorithm is used to purify the feature points for accurate matching. Experiments show that compared with the traditional image matching method, the proposed algorithm improves the efficiency and accuracy of the UAV image registration technology.

1. Introduction
In 1999, David Lowe first proposed the classic Scale-Invariant Feature Transform (SIFT) algorithm, which was improved in 2004. The algorithm combines the scale space pyramid with the Gaussian kernel to improve the detection rate of feature points, which is very invariant to scale scaling, translation rotation, and illumination variation. However, the algorithm has problems such as complex description and long calculation time. Yan Ke et al. reduced the feature description sub-dimension and reduced the amount of computation by using the PCA-SIFT algorithm of principal component analysis. By improving the Random Sample Consensus (RANSAC) algorithm, Wu et al. proposed to use the fast iterative strategy to filter the feature point pairs to obtain more accurate transformation parameters, and achieve accurate matching of remote sensing images. The image matching algorithm based on improved SIFT proposed in this paper can achieve automatic image registration and improve the efficiency and accuracy of UAV image registration technology.

2. SIFT algorithm
The basic idea of the SIFT algorithm is that the execution of the SIFT method is divided into two parts: first, the feature points are identified for the image, and then the feature points are described. The algorithm detects a stable feature point in an image with different degrees of blur, and then constructs a multi-dimensional descriptor by using some information retained in the detection process to express the feature point. Specific steps are as follows:
(1) Scale space construction: Search for images in all scale spaces, and identify potential points of interest and rotation invariant through Gaussian differential functions. The construction of the scale space is shown in the following formula:
\[ L(x, y, \sigma) = G(x, y, \sigma) \ast I(x, y) \quad (1) \]
\[ G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2) \]

The \( L(x, y, \sigma) \) of equation (2) is a scale space image at a certain scale, which can be obtained by convolving the original image \( I(x, y) \) with a Gaussian kernel \( G(x, y, \sigma) \).

(2) Construction of Gaussian pyramid: DOG (difference of guassians) pyramid is defined for the difference of pixel points in statistical images at different scales:
\[ D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) \times I(x, y) \]
\[ = L(x, y, k\sigma) - L(x, y, \sigma) \quad (3) \]

(3) Detection of spatial extremum: The key point is the local extremum point with direction information detected under images of different scale spaces. These points will not disappear due to changes in lighting conditions, such as corner points, edge points, bright spots of dark areas, and dark spots of bright areas. Since the two images have the same scene, they respectively extract their respective stable points. There will be a matching point between them.

(4) Construction of SIFT descriptor: After the above steps, the feature points existing at different scales have been found. In order to achieve image rotation invariance, the direction of the feature points needs to be assigned. The gradient distribution characteristics of the neighborhood pixels of the feature points are used to determine the direction parameters, and then the gradient histogram of the image is used to obtain the stable direction of the local structure of the key points.

3. Improved SIFT UAV image matching algorithm
The accuracy of the UAV image registration algorithm is low, and the SIFT algorithm feature descriptor is computationally complex and the time efficiency is low. This paper uses the Shi-Tomasi corner to detect the feature point operator and uses the SIFT feature vector to extract the feature points, then uses PCA to reduce the dimension of the feature vector. The specific algorithm flow chart is shown in Figure 1. The algorithm improves the time efficiency while ensuring high accuracy and good stability.

**Figure 1. Algorithm flow chart**

3.1. Shi-Tomasi detection feature points
The Shi-Tomasi algorithm was proposed by J. Shi and J. Tomasi in 1994. This method is an improvement to the Harris algorithm. In the Harris algorithm, it is determined whether or not a corner point is based on a combination of two eigenvalues of the covariance matrix M. In the Shi-Tomasi algorithm, it is determined whether the corner point is based on whether the smaller eigenvalue is greater than the threshold. This judgment is based on the fact that a smaller eigenvalue indicates that the variance in the direction of the eigenvalue is smaller, and if the smaller variance can be greater than the threshold, the change in this direction satisfies the judgment requirement of the corner point.
3.2. Feature point extraction

For the extraction of feature points, this paper uses the improved SIFT algorithm to experiment $D(x,y,\sigma)$ to improve the Gaussian pyramid difference function, and perform three binary Taylor expansion to achieve sub-pixel precision.

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X$$ (4)

In equation (4) $X = (x,y,\sigma)$, in order to enhance the stability of the key points, the key points with an extreme value less than 0.03 are filtered out.

Since the Gaussian difference function has a strong edge response, the Gaussian difference function has a large curvature along the edge, and the vertical edge has a small curvature. The eigenvalue of the matrix is proportional to the curvature of the Gaussian difference function, then dropping the unstable key points near the edge.

In order to ensure the rotation invariance of the feature vector, the coordinate axis is rotated by $\theta$ angle in the nearby neighborhood, that is, the coordinate axis is rotated into the main direction of the feature point, and the new coordinates of the pixels in the neighborhood after the rotation are:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$ (5)

3.3. Feature matching

After extracting the feature points according to the above analysis, the BBF algorithm can be used to first perform coarse matching on the feature points, and then the RANSAC algorithm is used to eliminate the mismatched points generated in the rough matching process to achieve accurate matching of the images.

BBF (Best Bin First) search algorithm: The BBF algorithm is a search algorithm developed on the basis of the K-D tree structure, and performs better in high-dimensional features than the K-D tree search algorithm. The BBF algorithm is an improved method of the K-D tree. It can be seen from the above-mentioned standard K-D tree query process that the "backtracking" in the search process is determined by the "query path", and does not consider some of the data points themselves on the query path. A simple improvement is to sort the nodes on the "query path". The BBF algorithm is an approximation algorithm proposed for high-dimensional data. This algorithm can ensure the priority of retrieving the space with the most likely neighbors. Therefore, we choose to use K-D tree to organize the feature points and use the BBF algorithm to search for feature points.

The RANSAC algorithm adopts the sampling and verification strategy. The main idea is to solve the eight parameters of the homography matrix $H$ that can be satisfied by one sample point. This is an important part of removing the wrong matching point, $(x,y)$ and $(x',y')$ are set to be matched Point coordinates. While the homography matrix $H$ describes the transformation relationship between the corresponding two sets of coordinates, that is, the geometric transformation relationship between images:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$ (6)

4. Experiment and analysis of results

In order to verify the effectiveness of the algorithm, this paper uses the aerial image of the crater in Bergen, a city in southwestern Norway, to experiment and compare the experimental results with the results of the SIFT algorithm. The experimental equipment parameters are: processor Intel Core 2 Duo T5750 2.00GHZ, memory 2GB) (DDR2667MHz). The software platform used in the experiment is:
Microsoft Visual C++ 2017 Express Edition, OpenCV 3.4.7, cmake-3.11.1-win64-x64. This algorithm is implemented using C++ language programming.

Figure 2. Raw image data

Figure 2 shows the original image data. Figure 3 shows the comparison of the feature point extraction results. The experimental process is to extract the key points of the two images first; then extract the SIFT descriptors from the four gradient patterns of each image; the gradient values of each key point and its surrounding points are sampled as descriptor vector elements, and the dimension of the feature vector is reduced by PCA. Figure 3 is a graph of the results obtained after the key points are extracted. The key points are marked by small circles.

Figure 3. Comparison of feature point extraction results

Figure 4. SIFT algorithm (left) and the algorithm matching and running results

Figure 4 shows the SIFT algorithm (left) matching and running results of the algorithm. We compare the two algorithms and calculate the matching time, the detected matching point pairs, the correct matching point pairs, and the matching rate.

It can be seen from Table 1 that the algorithm uses the Shi-Tomasi algorithm to calculate the feature corners only, and the detection time consumption is short. The improved SIFT algorithm proposed in this paper extracts more feature points than the traditional SIFT algorithm. The number is larger than the number of matching of the traditional SIFT algorithm, which shows the efficiency of
the proposed method in extracting feature points. The matching rate of the above experiment shows that the matching accuracy of the algorithm is relatively high.

|                | Match time | Detected match point pair | Correct match point pairs | Match rate |
|----------------|------------|---------------------------|---------------------------|------------|
| SIFT algorithm | 9.56 s     | 216                       | 156                       | 56.3%      |
| This paper algorithm | 3.12 s | 828                       | 549                       | 69.6%      |

5. Conclusion
In this paper, different strategies are adopted in different stages of the matching algorithm, and it shortens the running time. Using Shi-Tomasi corner detection and SIFT algorithm, RANSAC algorithm ensures the calculation accuracy. The algorithm makes full use of the gray information and corner position of the image. Experiments show that the proposed algorithm can achieve accurate and fast matching for images with large chromatic aberration and deformation, and its accuracy and speed are better than traditional matching algorithms, and increases the robustness of the algorithm.

Acknowledgement
Fund Project: Tianjin Science and Technology Commissioner Project (19JCTPJC54800); Tianjin University of Technology and Education Graduate Innovation Fund (YC19-30).

References
[1] ZHANG YIMING, FU HUI. A method of splicing aerialimage sequences of UAV forest land based on gap degree [J]. Journal of Beijing Forestry University, 2017, 39(6):107-115. (In Chinese)
[2] JIA YINJiang, XU ZHENAN, SU ZHONGBIN, et al. Remote sensing crop image mosaicing based on opti-mized SIFT algorithm [J]. Journal of Agricultural Engineering, 2017, 33(10):123-129. (In Chinese)
[3] QIN XUJIA,WANG QI, WANG HUILING,et al. Sequential remote sensing image mosaic algorithm based on opti-mal sutures[J]. Computer Science, 2015, 42 (10):306-310. (In Chinese)
[4] DENG TAO, HOU DONGPING, LE QINLAN, et al. Im-age mosaic method of small drone combined with SIFT and Harris algorithm [J]. Journal of Information Engineering University, 2015, 16(3):321-326. (In Chinese)
[5] Li Yan. Based on the improved SIFT algorithm of block model, the image matching method of UAV is studied [D]. Guiyang: Guizhou Normal University, 2016.(In Chinese)
[6] Feng Wenbin, Liu Baohua. Improved SIFT algorithm image matching research [J]. Computer Engineering and Applications, 2018, 54(3): 200-205. (In Chinese)
[7] Han Y, Choi J, Byun Y, et al. Parameter optimization for the extraction of matching points between high-resolution multisensors images in urban areas [J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(9):5612 – 5621.
[8] Cai G R, Jodoin P M, Li S Z, et al. Perspective-SIFT: An efficient tool for low-altitude remote sensing image registration [J]. Signal Processing, 2013, 93(11):3088-3110.
[9] BEKELE D, TEUTSCH M, SCHUCHERT T. Evaluationof binary keypoint descriptors[C]//IEEE International Conference on Image Processing. Melbourne, Australia: IEEE, 2013:3652-3656.
[10] Yan Ke, Sukthankar R. PCA-SIFT: A more distinctive representation for local image descriptors [J]. 2004, 2(2): 506－513.