An Image Matching Method Based on SIFT Feature Extraction and FLANN Search Algorithm Improvement

Shigang Wang*, Zhenjin Guo and Yang Liu
School of Electrical and Information Engineering, Guangxi University of Science and Technology, Liuzhou, 545006, China

*Corresponding author e-mail: 100001772@gxust.edu.cn

Abstract. In order to solve the problems of less feature information and high mismatching rate in traditional image matching algorithms, this paper proposed to extract and describe features based on the SIFT algorithm. FLANN algorithm was used to pre-match feature points, and RANSAC algorithm was used to optimize the matching results, so as to realize real-time image matching and recognition. Experimental results show that the proposed algorithm has better accuracy and better matching effect than traditional image matching methods.

Keywords: Image Matching Method, SIFT Feature Extraction, FLANN Search Algorithm

1. Introduction
Image matching refers to the method of finding similar images in two or more images through certain algorithms [1]. In the research process of digital image processing, image feature extraction and image matching have always been a key problem, playing a vital role in the fields of image registration, target detection, pattern recognition, computer vision and so on [2].

In 2004, Lowe published the SIFT algorithm, which extracted image feature points by constructing Gaussian scale space, finding extreme points, eliminating unstable feature points, determining the direction of key points and generating feature point descriptors [3]. SIFT algorithm introduced image pyramid structure, which reduces the computation amount without affecting the extraction of some feature points of the image. But the SIFT algorithm did not take into account the geometric constraints in the space information, resulting in a high mismatching rate, but also prone to appear particularly obvious mismatch and mismatch phenomenon.

Through the study of the above methods, this paper proposes a feature extraction and feature description based on SIFT algorithm, using FLANN algorithm to pre-match feature points, and using random sampling consistent RANSAC algorithm to optimize the matching results, so as to achieve real-time image matching and recognition.

2. SIFT Algorithm Principle
SIFT algorithm is effective for finding local features of image. The feature points, scales and direction descriptors are used to extract relevant features and match images. SIFT algorithm has a good effect in
feature extraction and detection. Changing the rotation angle and brightness does not affect the scale.

2.1. The Scale Space Was Constructed To Detect the Extreme Point and Obtain the Scale Invariance

Since Koendetink found that there was only one way to scale transformation, and that was to use the Gaussian kernel, image features were extracted at different scales to achieve scale invariance. First build the Gaussian Pyramid. Then the Dog pyramid was built. Finally, extremum detection is carried out on the basis of DOG pyramid.

2.1.1. Building the Gaussian pyramid. According to the two-dimensional Gaussian filtering function, Gaussian filtering is performed on the image.

\[ G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \]  

(1)

Where \( \sigma \) is called variance.

The representation \((x, y)\) in different scale Spaces can be obtained by the Gaussian kernel convolution.

\[ L(x, y, \sigma) = G(x, y, \sigma) \times f(x, y) \]  

(2)

\( \sigma \) is the scale space factor.

If it gets lower, the image gets smoother. Large scale is a general feature of the image. Small scale is an image detail feature. Therefore, the size of the scale factor is very important for the scale space description.

The Gaussian pyramid is built according to the scale function. The first layer of the first order of the Gaussian pyramid is the original image. The Gaussian pyramid has \( o \) and \( s \) levels. The ratio of two adjacent levels in the same order is \( k \). For example, if the ratio between level 1 and level 2 is \( \sigma \), then the ratio between level 1 and level 3 is \( k\sigma \). And each order image is one half the size of the previous order image.

Figure 1 with the Gaussian pyramid on the left and the Dog pyramid on the right, shows how to get the Dog pyramid from the Gaussian pyramid.

![Gaussian pyramid and Dog pyramid](image)

**Figure 1.** Gaussian pyramid and the Dog pyramid

2.1.2. Build the Dog Gauss pyramid based on the Gauss pyramid. A layer of Dog Gaussian pyramid can be obtained by the difference of two adjacent scale space functions of the same order.

The expression for Dog is defined as (3).
\[ D(x, y, \sigma) = (\tilde{g}(x, y, k\sigma) - \tilde{g}(x, y, \sigma)) \times I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \] (3)

2.1.3. The extremum points are detected and the key points are obtained. The 26 pixel values adjacent to each pixel in the Dog image need to be compared. The adjacent pixel points need to be taken from the middle layer on the same order (except for the bottom and top layer on the same order). Through comparison, the maximum or minimum value of the point is needed for the selected point, and then the key points are determined by eliminating the selected points and edge points with low contrast and noise sensitivity.

2.2. Determine the Direction Value of Key Points
The SIFT algorithm reduces computational requirements to the extent that a large number of feature description points are retained in the image by using the pyramid image structure. The formula is as (4) and (5).

\[ \alpha(x, y) = \sqrt{\left(L(x + 1, y) - L(x - 1, y)\right)^2 + \left(L(x, y + 1) - L(x, y - 1)\right)^2} \] (4)

\[ \theta(x, y) = \tan^{-1}\left(\frac{L(x + 1, y) - L(x - 1, y)}{L(x, y + 1) - L(x, y - 1)}\right) \] (5)

The adjacent pixel values determine the gradient direction of the key points. \( L \) is the spatial scale function of the key point, and \((x, y)\) is its adjacent pixel region of the key point. Gradient histogram is used to calculate the gradient direction of the adjacent pixel region. The horizontal axis of the gradient histogram represents the value of the gradient direction of the adjacent pixel region. The magnitude of the vertical axis is the neighborhood pixel gradient value. The horizontal axis of the gradient histogram ranges from \(0^\circ\) to \(360^\circ\). Every \(10^\circ\) is one unit. There are 36 units. The main peak of the histogram of the gradient direction represents the main direction of the key point [4]. If there are other peaks equal to 80% of the main peak value, the secondary direction of the key point. The main peak direction and multiple secondary peak directions determine the direction of the key points. Thus, the influence of image rotation on key feature points can be greatly reduced.

2.3. Generate Feature Vector Descriptors
To further describe the information of the key points. Therefore, it is important to determine the size of the neighborhood range of the key points, as shown in Figure 2, the smaller square represents the key neighborhood pixel. The arrow in the small square is the gradient direction of the neighborhood pixel, and the size of the arrow represents the gradient size. The small square in the upper left corner of figure (b) in Figure 2 consists of four small square in the upper left half of the image (a) in Figure 2. That is, (b) the direction of each square in the figure is the cumulative value of the direction of the squares in the Figure (a). Figure 2 shows (a) an \(8 \times 8\) neighborhood range and (b) a \(2 \times 2\) seed point (one seed per square). In order to enhance the anti-noise ability and the robustness of matching. We usually set the neighborhood to \(16 \times 16\). So you're going to have a \(4 \times 4\) seed point. In this way, the information of each key point is contained in the \(4 \times 4 \times 8 = 128\), dimensional feature vector.

3. FLANN Feature is Close to Pixel Matching
Muja and Lowe proposed FLANN algorithm based on KD tree operation or K-means tree in 2009. The effective search type and retrieval parameters are determined by the known data set distribution characteristics and the required spatial resource consumption [5,6]. The feature space required by FLANN algorithm is usually n-dimensional real vector space. The key is to search the nearest point from the nearest neighborhood of the instance point by Euclidean distance. The mathematical definition of Euclidean distance is shown as (6).

\[ D(x, y) = \|x, y\| = \sqrt{\sum_{i=1}^{d} (x_i - y_i)^2} \] (6)
In this method, the data points in the n-dimensional spatial query brush are divided into several specific regions according to the KD-Tree method. Its function is to retrieve the nearest Euclidean distance to the query point in its neighborhood [7].

In the n-dimensional space brush, all the Euclidean distances are stored by the structure of KD-tree, so the nearest point to the reference point can be found conveniently and effectively. The search process of KD-tree structure is recursive from top to bottom. Firstly, the target point is separated from the split point according to the specific datum, and then they are compared to judge whether the target point is to the left or to the right of the datum. Cycle comparison with corresponding points one by one until the target point is successfully searched [8].

3.1. Calculation of Transformation Parameters

In this paper, we study the affine change relation of translation and rotation between images. After the matching of feature points, the homography transformation matrix can be obtained according to the matched pair of feature points, so as to further obtain the relative change parameters between images.

Let \((X, Y)\) and \((x, y)\) be any pair of matching point pairs of two images, then (7):

\[
\begin{bmatrix}
X \\
Y \\
1
\end{bmatrix} = \begin{bmatrix}
h_{11} & h_{12} & h_{13} \\
h_{21} & h_{22} & h_{23} \\
0 & 0 & 1
\end{bmatrix} \times \begin{bmatrix}
x \\
y \\
1
\end{bmatrix} = H \times \begin{bmatrix}
x \\
y \\
1
\end{bmatrix} \tag{7}
\]

According to its change relation, \(H\) matrix can be expressed as (8).

\[
H = \begin{bmatrix}
\cos \theta & -\sin \theta & \Delta x \\
\sin \theta & \cos \theta & \Delta y \\
0 & 0 & 1
\end{bmatrix} \tag{8}
\]

\(\theta\) for rotation angle, \(\Delta x, \Delta y\) translation of x direction and y direction respectively.

It is necessary to improve the calculation accuracy of \(H\), RANSAC algorithm is adopted in this paper to eliminate mismatched point pairs, so as to obtain more accurate transformation parameters [9-10]. RANSAC algorithm is to select three groups of matching points of two images to match the six parameters of \(H\) matrix. The remaining matching points are judged by \(H\)-matrix, and the matching inner and outer points are distinguished. At the same time, the number of inner points is recorded, and then the parameters of \(H\) are updated by the new inner points. When the number of inner points
reaches the maximum, the optimal parameters of \( H \) are calculated, and the optimal H-matrix is obtained to complete the frame selection of the target image.

4. Analysis of Experimental Results

This experiment uses OpenCV3.12 to achieve the above algorithm. The test image is a grayscale image with \( 512 \times 384 \) pixels.

Figure 3 shows the results of BF matching for the image features extracted by ORB algorithm. Figure 4 is the result of BF matching with the features extracted from SIFT. Figure 5 shows the results obtained after FLANN matching of the features extracted by the SIFT algorithm. Figure 6 is the result of matching by this article. According to the analysis of the experimental results, the feature information of the two images in Figure 3, Figure 4 and Figure 5 are extracted to different degrees and matched. However, the method extracted in this paper produces more feature information points, which can be identified and selected by a box.

5. Conclusion

This paper discusses an improved image matching method for SIFT feature extraction and FLANN search algorithm. Due to the high mismatch rate and poor matching effect of traditional image matching algorithm, this paper proposes feature extraction and feature description based on SIFT algorithm, pre matching through FLANN algorithm, and using RANSAC algorithm to eliminate the mismatch problem in the matching results. Finally, the image matching and recognition are realized. Through the experimental data and analysis, it can be seen that the algorithm proposed in this paper is effective and effective in image matching effect and accuracy.

Acknowledgments

This work was supported by Innovation and Entrepreneurship Training Program for College Students in Guangxi (Grant No. 201910594025) and the Undergraduate Teaching Reform Project of Guangxi Higher Education (Grant No. 2018JGZ126).

References

[1] B. Y. Tan. Image Feature Extraction and Matching. The Graduate, 2008.
[2] Q. Chang, B. Zhang and J. L. Shao. Feature Image Matching Method Based on SIFT and
RANSAC. Journal of East China University of Science and Technology: Natural Science, 2012, 38 (6), pp. 747.

[3] Lowe D G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 2004, 60 (2), pp. 91-110.

[4] Akenine-Moller T, Moller T, Haines E. Real-Time Rendering. A. K. Peters, Ltd., 2002.

[5] Muja M and Lowe D G. Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. Proceedings of IEEE Conference on Computer Vision Theory and Applications. Lisbon, Portugal: IEEE Computer Society, 2009, pp. 331.

[6] Y. D. Feng and Y. sun. Image Matching Algorithm Based on SURF Feature Extraction and FLANN Search. Journal of Graphics, 2015 (4), pp. 650.

[7] S. Y. Liu, C. Q. Yang and X. L. Wei. Application of Neighbor-point Fast Search Method in Chaos Recognition. Journal of Huazhong University of Science and Technology, 2012, 40 (11), pp. 89.

[8] J. T. Cui, W. G. Liu and L, H, Zhou. A multi-resolution and high-dimensional image feature matching algorithm. Acta Photonica Sinica, 2005 (1), pp. 138-141.

[9] Fischler M A and Bolles R C. Random Sample Consensus: A Paradigm for Model Fit with Applications to Image Analysis and Automated Cartography. Communications of the ACM, 1981, 24 (6), pp. 381.

[10] W. B. Wang, X. L. Bai and Q. Xu. Feature Image Matching Based on SURF and RANSAC. Journal of Harbin University of Science and Technology, 2018, 23 (1), pp. 117.