Indoor image matching algorithm based on improved SIFT algorithm

Libiao Zhang*, Yuqing Liu, Chengqi Zhang, Bin Zhang and Meiyi Ge
School of Information Science and Technology, Northeast Normal University, Changchun, China

*Corresponding author e-mail: lbzhang@nenu.edu.cn

Abstract. Aiming at the indoor positioning by using SIFT algorithm, because the structure information is more or similar, and there is no obvious identification, which will lead to low accuracy and poor applicability, this paper proposes an improved SIFT algorithm for image matching. Choosing Junction feature which is more suitable for indoor structure, and combining with the improved CPW energy function method to eliminate mismatching points. By testing the actual image to compare and analyze get the feasibility and validity of the new algorithm.

1. Introduction

The proportion of human's indoor life trajectory has exceeded 70% of the total, and most of consumption is carried out indoors. Among them, the method based on computer vision has broader scope of application, higher accuracy and stronger anti-interference.[1,2] In computer vision, SIFT[3,4] algorithm is always used for image matching. Aiming at the problems of traditional SIFT algorithm in indoor image matching, this paper proposes a new method.

In the indoor environment, there are always no obvious characteristics, most of these building structures have high similarity.[5] The traditional SIFT image matching method has a large error when recognizing these structures. Therefore, the method based on Junction feature is used to match and explaining in this paper. Besides that briefly introduce the method of eliminating mismatching points, and an improved CPW energy function method is proposed. This method has a high degree of integration with Junction features, and uses the idea of support rate to reduce mismatching rate.

2. SIFT Algorithm

In the process of image matching, we need to find "interesting" information, that is, the information we need. Each image has its own characteristics, and when shooting for the same object, the image will have the similar characteristics. In the process of shooting, image rotation, enlargement, reduction, distortion and even blurring may occur, which will result in image matching problems. SIFT is an early and relatively perfect image matching algorithm, and it’s also a hot spot of improvement for a large number of researchers in this direction. SIFT algorithm needs to find the key point in different scale space, which is point feature. Moreover, the SIFT algorithm is very stable and will not be affected by image translation, rotation, zooming, light and shadow transformation, affine transformation, occlusion and noise.

Lowe's SIFT algorithm[6,7] can be divided to the following steps:
1. Establishing scale space, searching the position of image on all scales, as well as identifying possible rotation and scale-invariant candidate points which is mainly using Gauss differential function.

2. Determining key points, removing unstable ones and retaining stable points.

3. Determining direction, according to local gradient direction, each key point is assigned to one or more directions, and getting a reference direction to achieve rotation invariance.

4. Each key point has its own position, direction and scale. A descriptor is established to represent these attributes. The local gradient direction of the image is measured within the selected key point range.

SIFT algorithm has high stability and is hard to be disturbed. Besides that, it has strong plasticity and can improve performance by modification. But the shortcomings of SIFT algorithm are also obvious. Sometimes, the feature points of the algorithm are insufficient that can result in the poor matching results. Beyond that, if the edge of the target is smooth, it can’t accurately capture the feature points. Finally, it’s lack of real-time.

3. Improved SIFT Algorithm

When in low-texture scenes, such as indoor building images, point features can’t fully express the local architectural structure of the image, and the number of features extracted is small, so the matching results are unreliable. Most of the existing technologies are using line segment structure, which will result the destruction of the integrity of line feature. In point-line combination feature extraction, they are extracted separately, which is not friendly to indoor structure.

3.1. Junction Feature

Junction structure[8] is a more comprehensive and suitable feature for interior building structure. Junction structure is a geometric structure composed of two intersecting edge lines in different directions and the angle between them. He combines point and line features perfectly. Junction features can be divided into three types: L, Y and X. For an image, one of the detected K-branch structures can be expressed as:

$$J = \{p, r, \{\theta_k\}_{k=1,\ldots,K}\}$$  

Among them, P denotes the location of the center point of the branch of Junction structure, r denotes the length of the shortest branch, \(\theta\) denotes the direction of the branch.

In the above-mentioned Junction structural features, the lengths of branches in all directions are the same, but in practice, the lengths of branches in all directions are different. The anisotropic K-branch Junction can be expressed as formula (2), as shown in Figure 1.

$$J = \{p_1, p_2, \ldots, p_K\}$$

$$p_k = p + r_k (\cos\theta_k, \sin\theta_k)^T$$

**Figure 1.** Junction Structure.

3.2. Improved SIFT Algorithm

In the registration of indoor structure images, the traditional SIFT algorithm uses the method of point feature extraction. In some simple indoor spaces, the effective features extracted by this method are limited, and the matching accuracy is poor. Junction feature is more suitable for indoor image matching because of its unique structural features. So when searching and identifying key points, Junction feature is chosen to replace point feature as key points for further matching.
After Junction feature is detected, in order to eliminate the influence of contrast change, the normalized gradient is calculated in the neighborhood of the pixel. According to formula 3-2, \( P \) is the central point of Junction feature of a \( K \)-branch. For one of the branches, the neighborhood range along \( \theta \) is defined from the middle point. If the difference between the direction of a point in the neighborhood and theta is less than the threshold value, the point set can be added to the point set. RANSAC algorithm is used to fit the point set into a straight line \( L \). The shortest distance to \( L \) is the new \( P \) point. Through the above methods, the complete local structure of Junction features can be obtained.

After matching Junction features, the preliminary matching results are obtained. The flow chart of the improved SIFT algorithm is shown in Figure 2.

![Flow chart of improved SIFT algorithm.](image)

### 3.3. Eliminating Mismatching Points

In the process of matching, mismatching will occur because of the similarity of features or image distortion. Eliminating mismatches becomes more important.

Fischer and Bolles first proposed RANSAC algorithm\[9\]. RANSAC algorithm has the highest utilization rate in eliminating mismatched points, and it is simple and effective. This algorithm is one of the estimation models. Unlike the common methods, it chooses as few points as possible to estimate, but the impact is as wide as possible.

The advantage of RANSAC algorithm is the robustness in estimating model parameters. The disadvantage is that when calculating parameters, it will iterate indefinitely, and if the number of iterations is limited, the result may be inaccurate or even wrong. It has the chance to obtain a trusted model, which is proportional to the number of times.\[10,11\]Most of the methods to eliminate mismatching points are based on the improved RANSAC algorithm. Because choosing Junction feature, the method of eliminating mismatches will select more higher fitness to Junction one.

The CPW framework first divides the target image into grids, and then according to the minimum value of CPW energy function, optimizing each grid area, in order to reduce the matching error points. The basic CPW energy function methods include point alignment, smoothing and global alignment. Point alignment constrains the alignment of matching points. Smoothing term is to maintain the similarity of each grid before and after optimization. Global alignment constrains the range of meshes except points and lines. However, because of the structural characteristics of Junction feature, the line alignment term is added to restrict the line feature while keeping the line feature as straight as possible. The model of CPW energy function can be expressed as:

\[
E = \alpha E_p + \beta E_l + \gamma E_g + \rho E_s
\]  

Among them, \( \alpha, \beta, \gamma, \rho \) are used to represent the weighting coefficients of corresponding items, respectively, corresponding point alignment items, line alignment items, global alignment items and smoothing items.

After eliminating the mismatched points initially, according to the idea that the matched points around the mismatched points will be less than those around the correct points, the support rate around the matched points can be obtained by combining the partitioned mesh, and some points with low support rate can be removed to improve the accuracy. As shown in Figure 3, there are a large number of correct matching points around the correct matching points, while there are few or no supporters around the wrong matching points, so the wrong matching points can be removed by the support rate. The flow chart of eliminating mismatching points is shown in Figure 4.
4. Image Matching Method Based on Junction

In order to provide a good indoor location method, an image matching method based on Junction is proposed to improve the traditional matching algorithm for indoor structural features.

In this method, the concept of anisotropic Junction feature is introduced. The basic structure of the Junction feature is shown in formula (2). When searching and identifying key points, Junction feature is chosen to replace point feature as key points for further matching. After Junction feature is detected, in order to eliminate the influence of contrast change, the normalized gradient is calculated in the neighborhood of the pixel, and the Junction feature with relatively complete local structure is obtained. Combining Junction feature with SIFT algorithm can get a matching result.

Then, dividing the target image into meshes and constrained by point alignment, line alignment, global alignment and smoothing terms. The energy function model of CPW is shown in formula (3), and the minimum value of energy function is calculated according to the formula. Then, according to the grid distribution to calculate the support rate. The support rate principle is that there are more correct points around the correct matching points and fewer points around the wrong matching points.

The image matching method based on Junction is combining the improved SIFT algorithm with the improved CPW energy function method. The specific flow chart is shown in Figure 5.

5. Experimental Result and Analysis

According to the description of each step of the image matching algorithm based on Junction in this paper, this section will show the experiments result. In order to ensure the accuracy and reliability of the experiment, the experimental images are a group of images collected from the real campus office building. The image database includes blurring, rotation and zooming. As a test graph, the image set can test the comprehensive performance of the algorithm in practical use. Some results are shown in Figure 6. Table 1 shows the experimental data of Junction-based image matching algorithm and other matching algorithms in the test set, and visualizes them with Figure 7, so it can be more intuitive to compare between Junction-based image matching algorithm and other matching algorithms.
Table 1. Percentage comparison table of algorithm accuracy.

| Method       | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  |
|--------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| My           | 98.23 | 89.66 | 92.31 | 80 | 83.33 | 93.33 | 92.06 | 94.2 | 98.35 | 82.22 |
| SIFT         | 62.14 | 68.21 | 66.63 | 51.8 | 66.72 | 25.78 | 91.95 | 76.25 | 44.55 | 66.89 |
| Affine-SIFT  | 95.74 | 76.25 | 83.43 | 75.83 | 78.32 | 89.47 | 98.48 | 95.85 | 56.67 | 65.4 |
| Hessian-SIFT | 92.31 | 38.24 | 72.73 | 78.1 | 60.55 | 94.44 | 95.38 | 91.13 | 50.79 | 79.17 |

Figure 6. Partial Algorithmic Contrast Graph.

From the comparison of the results in Fig.6, it can be seen that the matching accuracy of this method is higher than the traditional SIFT algorithm. In indoor environment, the simpler the structure
is, the more obvious the gap between traditional SIFT algorithm and SIFT algorithm is. Such as fig. 6(d) has a simple structure, the accuracy of SIFT algorithm is very low, and the image matching algorithm based on Junction can match accurately. In the part of eliminating mismatching points, the test results show that when the values of $\alpha$, $\beta$, $\gamma$, $\rho$ in the CPW energy function model are 1, 1, 0.01 and 0.01, the energy function is the smallest and the result is the best. In summary, we can see that the proposed image matching algorithm based on Junction can perform better in indoor environment.

6. Conclusions

Based on the method of computer vision and along with the progress of image processing technology, researchers have been constantly exploring and researching. Using the respective feature information of the image, the image with the highest matching degree can be found and the location information can be obtained. However, when the indoor space environment is similar, the accuracy of this method will be greatly reduced.

In this paper, SIFT algorithm and common methods of eliminating mismatching points are introduced in detail. Besides that, this paper analyzes their advantages, disadvantages and application scope. By demonstration, an image matching method based on Junction feature is proposed. This method introduces the concept of Junction feature. And it explains the theoretical knowledge and structural characteristics of Junction. Then, the method of eliminating mismatched points is briefly introduced. According to Junction feature, an improved CPW energy function method is introduced to eliminate mismatched points. The improved CPW energy function method has a high degree of integration with Junction feature, and the mismatch rate can be reduced by using the idea of popularity.

After the actual environment image test, the accuracy of matching results is significantly higher than other traditional methods.

7. Acknowledgments

This work was financially supported by "the Fundamental Research Funds for the Central Universities".

References

[1] Chen Y, Chen R, Liu M, et al. Indoor Visual Positioning Aided by CNN-Based Image Retrieval: Training-Free, 3D Modeling-Free[J]. Sensors, 2018, 18(8):2692.
[2] Araújo P, Miranda R, Carmo D, et al. Air-SSLAM: A Visual Stereo Indoor SLAM for Aerial Quadrotors[J]. IEEE Geoscience & Remote Sensing Letters, 2017, PP(99):1-5.
[3] Liu J, Fu W, Wang W, et al. Image matching based on improved SIFT algorithm[J]. Chinese Journal of Scientific Instrument, 2013, 34(5):1107-1112.
[4] Hyunsup Yoon, Hwan-Ik Chung, Hernsoo Hahn. SURF algorithm with color and global characteristics[P]. ICCAS-SICE, 2009,2009.
[5] Ijaz, F., Hee Kwon Yang, Ahmad, A.W., Chankil Lee. Indoor positioning: A review of indoor ultrasonic positioning systems[P]. 2013.
[6] Dou J, Qin Q, Tu Z. Robust image matching based on the information of SIFT[J]. Optik, 2018, 171:S0030402618309021-.
[7] Sun J, Zhao Y, Wang S. Improvement of SIFT Feature Matching Algorithm Based on Image Gradient Information Enhancement[J]. Journal of Jilin University, 2018, 43(4):335-342.
[8] Xue N, Xia G S, Bai X, et al. Anisotropic-Scale Junction Detection and Matching for Indoor Images[J]. IEEE Transactions on Image Processing, 2017, PP(99):1-1.
[9] Chenchi Luo, McClellan, J.H.. Robust geolocation estimation using adaptive RANSAC algorithm[P]. Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on,2010.
[10] Tran N T, Tan D K L, Doan A D, et al. On-Device Scalable Image-Based Localization via Prioritized Cascade Search and Fast One-Many RANSAC[J]. IEEE Transactions on Image Processing, 2019, 28(4):1675-1690.
[11] Sengupta A, Elanattil S. New Feature Detection Mechanism for Extended Kalman Filter Based Monocular SLAM with 1-Point RANSAC[J]. 2018.