Research on image feature matching algorithm based on feature optical flow and corner feature

Tao Song¹, Bing Chen¹, Fu M. Zhao¹, Zheng Huang¹, Mu J. Huang¹
¹Chongqing Key Laboratory of Modern Photoelectric Detection Technology and Instrument, Chongqing University of Technology, Chongqing 400054, People’s Republic of China
E-mail: 52160745105@2016.cqut.edu.cn

Abstract: Image feature matching is an important part of visual odometry. In order to improve the accuracy of feature point matching for visual odometry, an accurate algorithm combining the pyramid feature optical flow and corner features based on image feature matching is proposed. Firstly, the Oriented Features from Accelerated Segment Test and Rotated BRIEF (ORB) algorithm is used to extract the image feature points quickly. Secondly, the local feature window is utilized to calculate the displacement vector of the image feature points and the pyramid Lucas–Kanade feature optical flow is used to track feature points. Then, to solve the problem of matching alignment and feature loss, the K nearest neighbour radius search is used as feature filter and speed-up robust feature matching (SURF) algorithm is introduced to eliminate redundant mismatch points and improved the matching rate. The comparison of experimental data shows that the proposed algorithm can get high matching rate. Compared with the traditional ORB feature matching algorithm, the algorithm had a significant improvement in real time and image feature matching accuracy.

1 Introduction

With the rapid development and application of the driverless, autonomous mobile robot and other technologies in the world, simultaneous localisation and mapping (SLAM) technology has been widely studied in academic circles [1–3], while visual SLAM (VSLAM) has become the most potential and future development direction due to its abundant information and low cost [4, 5]. VSLAM includes image matching, optimal filtering, loop detection, 3D reconstruction and so on. Compared with laser SLAM, VSLAM is difficult to be processed in real time because of its large computation and long-time consumption, and there are some problems such as image error recognition and so on, which limits the application of VSLAM technology [6]. Among the many research problems of VSLAM, image matching is the most basic and the most time-consuming part of the whole process, which has a significant impact on the VSLAM results. The accuracy and real time of image matching is directly related to the application of VSLAM technology and is very important to VSLAM.

Image matching mainly includes based on grey correlation [7], feature point extraction and relational matching [8]. Based on grey correlation method, it is greatly influenced by illumination, angle of view and scale, and its robustness is poor. Based on relational matching need to establish semantic network which is the application of artificial intelligence technology in image processing and needs powerful and complex hardware platform as the support, but there is no breakthrough at present. Image-matching algorithm based on feature points is the main direction of current image processing technology. After the scale invariant feature transform (SIFT) [9, 10] and speed-up robust feature (SURF) [11] algorithms were proposed, Rublee et al. [12] proposed a highly efficient ORB algorithm. In ORB algorithm, Features from Accelerated Segment Test (FAST) algorithm was used to detect feature points, and Binary Robust Independent Elementary Features (BRIEF) algorithm was used to calculate descriptor of feature points. The speed of this algorithm was 100 times faster than that of SIFT algorithm and 10 times faster than that of SURF algorithm. Since the high efficiency of ORB feature extraction algorithm, it had been applied to image processing very quickly [13–15]. VSLAM algorithm based on ORB features was also proposed by Mur-Artal, which made real-time operation of VSLAM became feasible [16–18].

Although the ORB algorithm had made a breakthrough in the efficiency of feature extraction, compared with the SIFT algorithm for rotation, affine and scale invariance, its anti-jamming and matching accuracy were reduced to a certain extent. For this reason, some scholars put forward some improved algorithms based on ORB algorithm. For example, to solve the problem of poor affine property of ORB algorithm, Hou et al. [19] proposed a fast affine invariant image matching algorithm based on ORB algorithm (AORB). Compared with the Affine-SIFT (ASIFT) algorithm [20], the computational speed was more than six times higher. Zeng et al. [21] proposed a basis for the low efficiency of ASIFT algorithm. The fast large angle image matching algorithm proposed by ORB improves the matching efficiency and was 10 times more efficient than the ASIFT algorithm. Lile et al. [22] reduced the spatial feature point search region by extracting ORB feature points combined with polar constraint, uniqueness constraint and order constraint. KNN was used as a matching strategy, even if the image was not proofread, it had a high matching rate.

As has been shown in [12], the ORB algorithm did not attempt to solve the scale invariance problem of BRIEF descriptor, but we can propose tracking or heuristic strategy to solve. Aiming at the problems of the scale change of target, the dynamic change of illumination and the incomplete occlusion of target, Lu and Li [23] proposed a feature point matching tracking method combining with bi-directional optical flow constraint. The algorithm could effectively solve the problems of target deformation and incomplete occlusion in complex scenes, and also had strong robustness to scale and affine variation.

In this paper, we proposes an image matching method that the ORB corner feature is tracked with the characteristic optical flow, which can optimise point features. By using standard binocular datasets and real image data, the experimental results show that the proposed algorithm has higher matching efficiency, better real-time performance and better robustness than the traditional matching algorithm.

The rest of the paper is organised as follows: Section 2 describes the related work. Section 3 presents the mathematical details of the proposed algorithm and Lucas–Kanade optical flow and ORB (LK-ORB) matching system. In Section 4, we compare
In 1981, Lucas and Kanade [26] proposed an LK optical flow algorithm for dense optical flow. However, the LK algorithm has a disadvantage that is when the object moves faster, the optical flow error is larger. In order to reduce the moving displacement of the moving object in image, we can reduce the size of image. Therefore, Bouguet proposed an improved LK algorithm based on pyramid using affine transformation [27], which uses a coarse-to-fine strategy for optical flow estimation, and the improved algorithm is especially effective in tracking feature points. The LK optical flow algorithm based on Gauss pyramid is shown in Fig. 1, and the tracking result of pyramid LK optical flow is shown in Fig. 2.

2.2 Feature extraction algorithm based on ORB

In 2011, Rublee et al. [12] proposed ORB feature matching algorithm, which adopts the characteristics of the FAST and the BRIEF. The problem that the BRIEF algorithm has no rotation invariance and sensitivity to noise, and the problem that the FAST feature points have no direction is solved. The FAST feature point detection schematics is shown in Fig. 3. FAST is a fast corner feature extraction algorithm which detects the obvious change of local pixel greyscale. The simple description is that the grey value of a pixel in the image is larger or smaller than that of the surrounding pixel, we believe the point may be a feature point. However, the FAST corner lacks rotation invariance. In order to make the FAST corner directional, the ORB algorithm uses greyscale centroid method in the main direction of FAST to make it directional. The concrete method is that taking the feature point as the center, and calculating the centroid position in the neighbourhood $N$, then constructing the vector with the feature point as the starting point and the centroid as the end point, so the direction of the vector is the direction of the feature point. The neighbourhood moment is defined as

$$M_{p,q} = \sum_{(x,y) \in N} x^p y^q I(x,y)$$

where $I(x,y)$ represents grey value of the image. The position of the regional centroid is

$$C = \begin{pmatrix} M_{0,0} & M_{0,1} \\ M_{1,0} & M_{1,1} \end{pmatrix}$$

The angle between the centroid and the feature point is the main direction $\theta$ of the FAST feature point

$$\theta = \arctan(M_{1,1}, M_{1,0})$$

The ORB algorithm uses the binary BRIEF descriptor to describe the local feature of the extracted feature points, and improves the shortcomings of the BRIEF descriptor to form the rBRIEF descriptor. The main ideas are as follows. Firstly, smoothing the image block. Then randomly selecting $N$ sets of point pairs according to the Gaussian distribution near the feature points (general the $N$ equal to 128, 256, 512 etc.). In the paper, we select the $N$ equal to 256, and define a binary bit for each point pair defined as

$$r(p, x, y) = \begin{cases} 1 & \text{if } p(x) < p(y) \\ 0 & \text{otherwise} \end{cases}$$

where $p(x)$ and $p(y)$ represent the pixel grey value of random point pair. The random point pair is processed in turn to generate a unique $N$-dimension binary string, that is called the BRIEF descriptor. We can express it as

$$f_N(p) = \sum_{1 \leq i \leq N} 2^{i-1} r(p, x_i, y_i)$$

Since the BRIEF does not have rotation invariance at this time, BRIEF has been improved in the ORB. Defining $N$ sets of point pairs to form the matrix $S$ for initializing BRIEF.
The optical flow method is a fast and high frequency algorithm, which has high precision for the scene with small displacement between frames. ORB algorithm is a high accuracy, feature extraction speed, but low frequency algorithm. The two algorithms have good performance complementarity, so we proposed a fusion matching algorithm of pyramid LK-ORB. The flowchart of our method is shown in Fig. 4.

The LK-ORB fusion algorithm includes the following steps:

- Firstly, we get two successive frames \( I_1 \) and \( I_2 \) in image sequence and preprocess the two image by greyscale. Then extract the ORB feature and the key points separately. The grey value of a pixel in the image is larger or smaller than that of the surrounding pixel, so the point may be a feature point. In this paper, ORB features are extracted from 8-level image pyramids with a scale factor of 1.2. The sum of greyscale differences in the neighbourhood of ORB feature points is as follows:

\[
N = \sum_{x \in N(p)} |I(x) - I(p)| > \varepsilon_d
\]

where \( I(x) \) represents the grey level of any point in the neighbourhood of \( p \) in Fig. 3. \( I(p) \) denotes the grey level of the neighbourhood center, \( \varepsilon_d \) is the threshold of grey difference. If \( N \) is greater than a given threshold (generally \( 3/4 \) of the point in the adjacent domain), then \( p \) is considered to be a characteristic point.

- The feature points of frame \( I_1 \) and \( I_2 \) are tracked by using pyramid Lucas–Kanade optical flow tracking algorithm. During this period, the threshold \( T \) is set to \( N/2 \). When the number of tracking points is less than this threshold, only a few feature points are tracked, the current frame feature points are re-extracted or the current frame is replaced by the previous frame optical flow tracking.

- Calculate the sum of squared intensity difference of feature points as the basis for measuring the quality of tracking in the image window \( W \). For the greyscale images collected in the last step, assume that a feature window \( W \) containing information about ORB feature points. Record the image frame of the time \( t \) and \( t + \tau \) as \( I(x,y,t) \) and \( I(x,y,t + \tau) \), respectively, and satisfied

\[
I(x - dx, y - dy, t) = I(x, y, t + \tau)
\]

That is, each pixel of the frame \( I(x,y,t + \tau) \) can be determined by the pixel translation amount \( d = (dx, dy) \) of the corresponding feature window of frame \( I(x,y,t) \). Defined \( \epsilon(d) \) as

\[
\epsilon(d) = \int \int_W [I(x + dx, y + dy) - I(x, y)]^2 ;
\]

where \( W \) represents a given feature window, \( \omega(x, y) \) denotes a window weighting function, its value is generally set to 1, or

\[
S = \begin{bmatrix}
x_1 & x_2 & \ldots & x_N \\
y_1 & y_2 & \ldots & y_N \end{bmatrix}
\]

The affine transformation matrix \( R_0 \) is determined by the main direction \( \theta \) detected by the FAST corner point, and the description matrix \( S_0 \) is obtained by affine transformation of matrix \( S \)

\[
S_0 = R_0S = \begin{bmatrix}
\cos \theta & \sin \theta & 0 \\
-\sin \theta & \cos \theta & 0 \end{bmatrix} \begin{bmatrix}
x_1 & x_2 & \ldots & x_N \\
y_1 & y_2 & \ldots & y_N \end{bmatrix}
\]

The rBRIEF descriptor with rotation invariance is obtained

\[
g_{\theta}(x,y) := f_{\theta}(x,y) \in S_0
\]

3 Methodology

Gauss distribution function. In this paper, we use Gauss distribution function to emphasise the central region of the feature window. In order to minimise \( \epsilon(d) \), we expand \( \epsilon(d) \) with Taylor, and the first and second terms are reserved. Expressed the first-order coefficient as \( g \)

\[
g = \left[ \frac{\partial}{\partial x} \frac{I(x + d) + I(x - d)}{2}, \frac{\partial}{\partial y} \frac{I(x + d) + I(x - d)}{2} \right]^T
\]

Deriving on both sides of the transformed \( \epsilon(d) \). When the derivative equals to 0, we obtain the minimum value. Namely

\[
\frac{\partial \epsilon(d)}{d} \approx \int \int_W [I(x + dx, y + dy) - I(x, y)] + g^Td.
\]

For two successive greyscale images, solved (14) to obtain the displacement \( d \) of the feature points. Due to the introduction of errors when Taylor expansion, the Newton method is used to eliminate the error in the paper

\[
d_{k+1} = d_k + [\begin{bmatrix} R_1 & \text{g}_{dx} & \text{g}_{dy} \\ \text{g}_{dx} & \text{g}_{yy} & \text{g}_{y} \\ \text{g}_{dy} & \text{g}_{y} & \text{g}_{y} \end{bmatrix}]^{-1} \int \int_W [I(x + dx, y + dy) - I(x, y)]^T \omega(x,y)d(x,y)
\]

where the iterative initial value \( d_0 \) is set to 0. \( d_k \) represents the displacement value obtained by the \( k \)th iteration. In order to improve the speed and accuracy of the image feature matching, the termination iteration condition is adopted as the horizontal or the vertical displacement increment is <0.1.

Finally, in order to deal with the matching alignment problem of image features and find the missing features, we first use the KNN to match. The algorithm sets the ratio of the optimal match to the sub-optimal match to be <0.7, and sets \( K \) equal to 2. The KNN is used as a filter to remove the confusing matches (if the two feature points in the image frame are too close or the KNN
At the same time, there will be some unmatched points or mismatched points, we finally use the RANSAC method [28] to remove redundant matching points and obtain the optimal matching.

4 Results

4.1 Data set matching experiment of LK-ORB algorithm

In this paper, the proposed LK-ORB algorithm is used to match the 52 groups’ images of the Málaga datasets [29] and New College datasets [30], respectively. The resolution of image is unified to 640 × 480 pixels. The experimental platform is Windows 7 operating system, AMD A6 CPU, 4GB RAM, and the programming environment is Visual Studio 2010 of based on OpenCV 3.0. The result of feature matching is shown in Fig. 5.

In order to further verify the effectiveness and robustness of our algorithm, we use the RER-720P2CAM binocular camera to capture the images of three scenes in the indoor and outdoor environment, including the street (Street), Square (Square), Laboratory (Laboratory). In each scene, 52 sets of images are collected as the matching images, and the resolution of the images is 640 × 360. The results of feature matching experiments using ORB algorithm in three real datasets of Street, Square and Laboratory are shown in Fig. 6. The results of feature matching experiments by LK-ORB are shown in Fig. 7.

4.2 Analysis of experimental results

In this paper, the continuous 52-frame binocular sequence images are selected in the standard datasets and the actual data set for real time, feature matching and accuracy testing. The algorithm uses the first and second frames of the sequence images as the initialisation frame of feature extraction preprocessing. We set the total number of initial feature points extracted by ORB algorithm to be 500. The experimental results show the matching results of the LK-ORB algorithm in Table 1. It can be seen from the table that the optical flow tracking method simplifies the matching pairs and eliminates the mismatch points under the algorithm studied in this paper. The stability of the feature points after tracking is improved, and the average matching rate of feature tracking can reach >98%.

Comparing the matching results of ORB algorithm in Fig. 6 with the LK-ORB algorithm in Fig. 7, no matter where in indoor or outdoor scene, the matching result of ORB algorithm has mismatch, and the matching performance is not as good as the robustness of the algorithm proposed in this paper. From the experimental data in Table 1, we can see that the initial matching error is ∼28.2% when the LK optical flow tracking algorithm is not used, but the mismatch is eliminated by using our proposed algorithm and combining with RANSAC algorithm, the matching degree is guaranteed very well. The error is only ∼1.9%. All the experimental data show that the proposed algorithm has better results in real-time and robustness, and the overall performance of the algorithm is better than the ORB algorithm.
In order to improve the robustness and accuracy of image feature extraction in visual odometry, a fast feature extraction and matching algorithm is proposed that combines feature optical flow with corner features based on the ORB algorithm. Compared with the traditional ORB algorithm, our algorithm has better robustness in dynamic scene and complex environment, which can eliminate a large number of mismatches and achieve high accuracy image matching. Based on the extraction of feature points, the algorithm uses K nearest neighbour radius to search for feature points in local image regions and uses the tracking characteristics of optical flow to track feature points in the region, so that the extracted image feature points are more robust and then the RANSAC algorithm is used to eliminate a few redundant features. The proposed algorithm not only reduces the mismatch but also improves the robustness of the algorithm in the dynamic scene. Finally, we analyse and verify the proposed algorithm through multiple sets of contrast experiments, and proves that the algorithm is robust and superior to the feature matching of images.

### 5 Conclusion

In order to improve the robustness and accuracy of image feature extraction in visual odometry, a fast feature extraction and matching algorithm is proposed that combines feature optical flow with corner features based on the ORB algorithm. Compared with the traditional ORB algorithm, our algorithm has better robustness in dynamic scene and complex environment, which can eliminate a large number of mismatches and achieve high accuracy image matching. Based on the extraction of feature points, the algorithm uses K nearest neighbour radius to search for feature points in local image regions and uses the tracking characteristics of optical flow to track feature points in the region, so that the extracted image feature points are more robust and then the RANSAC algorithm is used to eliminate a few redundant features. The proposed algorithm not only reduces the mismatch but also improves the robustness of the algorithm in the dynamic scene. Finally, we analyse and verify the proposed algorithm through multiple sets of contrast experiments, and proves that the algorithm is robust and superior to the feature matching of images.

### 6 Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 61701056), Chongqing Natural Science Foundation (No. cstc2016jcyljA0308), and Postgraduate Innovation Fund Project of Chongqing University of Technology (No. ycx2018233). Finally, the authors are grateful to the anonymous reviewers for their helpful comments and constructive suggestions.

### 7 References

[1] Song, W., Fu, M., Yang, Y.: 'An efficient traffic signs recognition method for autonomous vehicle', *Robot*, 2015, 37, (1), pp. 102–111
[2] Cadena, C., Carlone, L., Carrillo, H., et al.: 'Past, present, and future of simultaneous localization and mapping: towards the robust-perception age', *IEEE Trans. Robot.*, 2016, 32, (6), pp. 1309–1332
[3] Hong, S., Kim, J., Pyo, J., et al.: ‘A robust loop-closure method for visual SLAM in unstructured seabed environments’, *Auton. Robots.*, 2016, 40, (6), pp. 1095–1109
[4] Zuo, X., Xie, X., Liu, Y., et al.: ‘Robust visual SLAM with point and line features’. IEEE Intl. Conf. on Intelligent Robots and Systems, Vancouver, Canada, 2017, pp. 1775–1782
[5] Karlsson, N., Bernardo, E.D., Ostrowski, J., et al.: ‘The vSLAM algorithm for robust localization and mapping’. IEEE Intl. Conf. on Robotics and Automation, Orlando, FL, USA, 2006, pp. 24–29
[6] David, V., Arturo, G., Reinoso, O., et al.: ‘Improved omnidirectional odometry for a view-based mapping approach’, *Sensors*, 2017, 17, (2), p. 325
[7] Ying, L.: ‘Improvement of image matching algorithm based on gray correlation’, *J. Appl. Opt.*, 2007, 11, (1), pp. 68–73
[8] Zhong, W.H., Ning, X.B., Wei, C.J.: ‘A fingerprint matching algorithm based on relative topological relationship among minutiae’. IEEE Intl. Conf. on Neural Networks and Signal Processing, Nanjing, China, 2008, pp. 225–228
[9] Lowe, D.G.: ‘Object recognition from local scale-invariant features’. ICCV IEEE Computer Society, Kerkyra, Corfu, Greece, 1999, p. 1150
[10] Lowe, D.G.: ‘Distinctive image features from scale-invariant keypoints’, *Int. J. Comput. Vis.*, 2004, 60, (2), pp. 91–110
[11] Bay, H., Tuytelaars, T., Gool, L.V.: ‘SURF: speeded up robust features’. Proc. of the 9th European Conf. on Computer Vision, Graz, Austria, 2006, pp. 404–417
[12] Rublee, E., Rabaud, V., Konolige, K., et al.: ‘ORB: an efficient alternative to SIFT or SURF’. 2011 IEEE Intl. Conf. on Computer Vision, Barcelona, Spain, 2012, pp. 2564–2571
[13] Wang, L., Qiu, Z.: ‘Combining color histogram and ORB features for robust visual tracking’. IEEE Eighth Int. Conf. on Natural Computation, Chongqing, Sichuan, People’s Republic of China, 2012, pp. 519–522
[14] Hu, Z., Li, B., Hu, Y.: ‘Fast sign recognition with weighted hybrid K-nearest neighbors based on holistic features from local feature descriptors’. *J. Comput. Civ. Eng.*, 2017, 31, (5), p. 04017034
[15] Zha, Y., Shen, X., Chen, H.: ‘Copy-move forgery detection based on scaled ORB’, *Multimedia Tools Appl.*, 2016, 75, (6), pp. 3221–3233
[16] Mur-Artal, R., Montiel, J.M.M., Tardos, J.D.: ‘ORB-SLAM: a versatile and accurate monocular SLAM system’, *IEEE Trans. Robot.*, 2015, 31, (5), pp. 1147–1163
[17] Wang, J., Zheng, S., Du, Y., et al.: ‘Study on the ORB algorithm in the application of monocular SLAM’, *Nanotechnology*, 2015, 2, (3), p. 186
[18] Mur-Artal, R., Tardos, J.D.: ‘ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras’. *IEEE Trans. Robot.*, 2016, 33, (5), pp. 1255–1262
[19] Yi, H., Shihlin, Z., Lin, L., et al.: ‘Fast fully affine invariant image matching based on ORB’, *Comput. Eng. Sci.*, 2014, 36, (2), pp. 303–310
[20] Brese, C.L., Ju, J.Z., Uy, B.: ‘An improved ASIFT algorithm for matching repeated patterns’. IEEE Int. Conf. on Image Processing, Hong Kong, People's Republic of China, 2010.

[21] Zeng, Q.H., Chen, Y., Wang, Y.S., et al.: ‘A fast matching method for large viewpoint changes images based on ORB algorithm’, Control Decis., 2017, 32, (12), pp. 2233–2239.

[22] Lile, H., Erwei, L., Zhaoxing, L., et al.: ‘Fast stereo matching algorithm based on ORB operator’, Comput. Eng. Appl., 2017, 53, (1), pp. 190–194.

[23] Lu, S.N., Li, X.H.: ‘Vehicle tracking method using feature point matching combined with bidirectional optical flow’, J. Transp. Syst. Eng. Inf. Technol., 2017, 17, (4), pp. 76–82.

[24] Hetherington, R.: ‘The perception of the visual world’ (Houghton Mifflin, Oxford, UK, 1950), p. 172.

[25] Bradski, G., Kaehler, A.: ‘Learning OpenCV’ (O'Reilly Media, Newton, MA, USA, 2014, 3rd edn.)

[26] Lucas, B.D., Kanade, T.: ‘An iterative image registration technique with an application to stereo vision’. Int. Joint Conf. on Artificial Intelligence, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1981, pp. 674–679.

[27] Bouguet, J.Y.: ‘Pyramidal implementation of the Lucas Kanade feature tracker description of the algorithm’, Opencv Doc., 1999, 22, (2), pp. 363–381.

[28] Fischler, M.A., Bolles, R.C.: ‘Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography’, Commun. ACM, 1981, 24, (6), pp. 381–395.

[29] Blanco-Claraco, J.L., Moreno-Duenas, F.A., Gonzalez-Jimenez, J.: ‘The Málaga urban dataset: high-rate stereo and LiDAR in a realistic urban scenario’, Int. J. Robot. Res., 2014, 33, (2), pp. 207–214.

[30] Smith, M., Baldwin, I., Churchill, W., et al.: ‘The new college vision and laser data set’, Int. J. Robot. Res., 2009, 28, (5), pp. 595–599.