IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 17, 2019, accepted October 29, 2019, date of publication November 4, 2019, date of current version November 14, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2950909

Multi-Stage Matching Approach for Mobile Platform Visual Imagery

YU CHEN[1], LI YAN[1], BO XU[1], AND YUXUAN LIU[2]
[1]School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China
[2]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

Corresponding authors: Yu Chen (chenyuphd@whu.edu.cn) and Li Yan (lyan@sgg.whu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China (NSFC) through the Data Fusion Processing Technology Study of Oblique Photography and Ground Movement Survey Networks Project under Grant 41771486.

ABSTRACT Unpredictable texture structure and motion blur continuously exist in mobile platform visual imagery and seriously reduce the similarity between images. Thus, accurate, stable, and well-distributed matches to follow the accurate pose estimation of the platform are difficult to obtain. To solve such problems, an effective image matching method for mobile platform visual imagery is presented in this study. The proposed method includes three steps, namely, standard initial matching, transformation matrices evaluation and matching propagation. Firstly, an oriented FAST and rotated BRIEF (ORB) method was used to obtain the number of matches and the initial projective transformation relationship between an image pair. Secondly, an evaluation function was set to choose the suitable rotation matrix for the image scene. Finally, geometric correspondence matching was utilized to propagate matches and produce additional reliable matching results. The geometric correspondence matching used the geometric relationship between the image pair and found more suitable matches than the standard ORB matching. Comprehensive experiments on TUM and ICL-NUIM dataset images showed that the proposed algorithm performs better in terms of correct matches, satisfactory matching rate, and higher matching accuracy than the standard ORB and ORB-slam2 initial match methods.

INDEX TERMS Mobile platform visual imagery, image matching, ORB, propagation matching, geometric correspondence matching.

I. INTRODUCTION

Image matching is the process of finding corresponding points on multi-view images of the same area [1]. This process is not only the premise of automatic image registration [2], [3] but also the basis for image sensors to acquire their own pose and for track trajectory, positioning [4]–[6], target recognition [7], and 3D reconstruction [8]. With the diversification and miniaturization of mobile platforms, the demand for fast and stable visual image matching algorithms is rapidly increasing.

Given the requirement for real-time performance, the corresponding points between the mobile platform visual images are difficult to match. This is because the algorithms require time to avoid the influences from the geometric distortion caused by distance transformation and viewing angle conversion. The algorithms also require time to avoid the photometric distortions, such as illumination and seasonal changes, and it's hard to find an effective way to avoid costing time. The standard oriented FAST and rotated BRIEF (ORB) feature matching algorithm [9] based on FAST corner detection [10], described by the binary robust independent elementary feature (BRIEF) descriptor [11], [12], and matched by Hamming distance can quickly obtain a matching result of two images without GPU acceleration given its high computational efficiency. Therefore, the ORB algorithm is widely used in systems, such as visual odometers which require high real-time algorithms, or hardware devices, such as mobile phones that have low energy efficiency. However, the ORB algorithm becomes prone to mismatch when the images acquired by the sensor is blurred due to the rapid motion of the mobile platform and overly monotonous or complicated texture structure in the scene. The matching quality of the image pairs will directly affect the autonomous navigation and modeling quality of the subsequent mobile platform. To improve the performance of the

The associate editor coordinating the review of this manuscript and approving it for publication was Tomasz Trzcinski.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

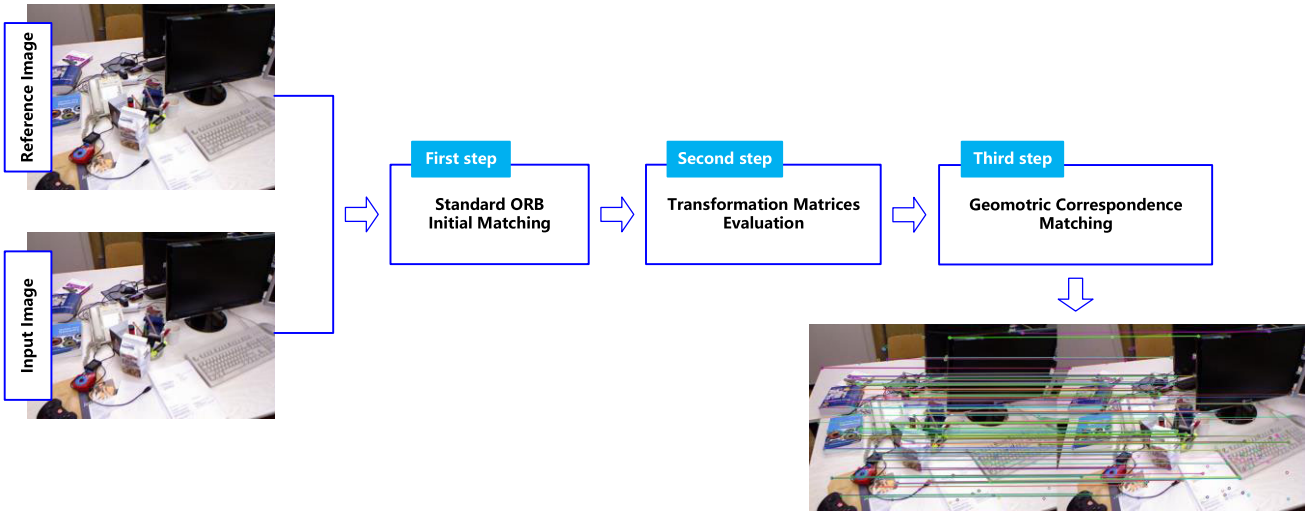


FIGURE 1. The framework of geometric correspondence - oriented FAST and rotated BRIEF (GC-ORB).

ORB algorithm, researchers have mainly applied the following approaches: (1) improving the feature point extraction strategy [7], [13]–[19], (2) enhancing the feature descriptor [20]–[22], and (3) upgrading the strategy of matching feature points [23]–[28]. For example, to improve the distribution uniformity of feature points, enhance the matching efficiency, and increase the correct matching rate, the ORB-slam2 [29] initial matching algorithm boosts the feature point matching efficiency; this improvement is achieved by rasterizing the entire image and modifying the minimum Hamming matching distance to 0.7–0.9 times the sub-optimal matching distance; moreover, the histogram statistics are used to describe the sub-rotation main direction angle, and only the top three direction feature points are selected in the statistics to improve the correct matching rate.

Scholars have proposed multiple enhancement matching strategies for multiple feature matching algorithms to improve the matching results of the feature points. Yu *et al.* [23] used the K-value nearest neighbor algorithm for coarse matching and applied the random sample consensus (RANSAC) algorithm for fine matching and removing the wrong matching point pairs. Xin *et al.* [24] enhanced the ORB algorithm through the progressive sample consensus (PROSAC) algorithm. In comparison with the traditional RANSAC algorithm, the PROSAC algorithm considers the correlation problem of sampling points and thus improves the robustness and efficiency of the ORB algorithm. Geng [30] proposed the use of the nearest six matches around the candidate points to calculate the affine transform in the local region to predict the match and then compute the normalized correlation coefficient (NCC) between the candidate point and the point around the predicted position. If the NCC of the point with the highest value is greater than the predetermined value, the points were identified as matching point pairs. Liu *et al.* [31] determined the homography matrix between two images in accordance with the existing UR-SIFT matching point pairs and obtained new exact matching points through geometric correspondence matching based on the homography matrix; in addition, these researchers applied probability relaxation to ascertain the correct point pairs which were unsuccessfully matched previously. The algorithm improved the matching accuracy in multi-source remote sensing image matching and significantly increased the number of correct matching point pairs.

Based on these previous works, the present study proposes an enhanced matching algorithm called geometric correspondence – oriented FAST and rotated BRIEF (GC-ORB) to be applied to mobile platform visual images. Firstly, the proposed algorithm extracts high-quality FAST key points at various scales using image pyramids and calculates the BRIEF descriptors. An initial matching is performed, and the rotation matrices between the image pair are obtained. Secondly, the obtained fundamental and homography matrices are evaluated by the evaluation function. Finally, the rotation matrix which is suitable for the image scene is selected for the geometric correspondence matching of all feature points. The selection aims to obtain as many correct matches as possible and to delete the mismatched points to make the final match. Figure 1 illustrates all the steps of the entire algorithm.

II. METHODOLOGY

The ORB is a feature extraction algorithm based on the FAST feature points and their BRIEF descriptors. The algorithm can acquire enough feature points from two frames of images efficiently and quickly match the homonymous points. However, the matching results of this algorithm often have mismatches. In order to solve the mismatching problem caused by image noise, dynamic blurring or texture with monotonous line features in ORB algorithm, this paper proposes a strategy to improve the matching quality of the algorithm by adding two steps: transformation matrix evaluation and geometric correspondence matching.

A. STANDARD ORB INITIAL MATCHING

1) FAST FEATURE EXTRACTION

This algorithm aims to obtain the numbers of relatively accurate features across the entire range of the image as soon as possible. The specific steps are presented as follows:

1) Extract the number of FAST feature points by roughly comparing the gray brightness difference around all the pixels. Use each point on the image (with the brightness similar to I_p) as a center, pick 16 pixels on a circle (with a radius of 3), and choose the point as suitable if 12 or more of the 16 pixels are lighter than $1.2I_p$ or darker than $0.8I_p$.
2) Remove local clustered feature points through non-maximum suppression, and only keep the points that can respond to the maxima in a certain area.
3) Determine the approximate total number N of the feature points. The expected value is typically set as 500. Calculate the Harris response values for the original corner points, and select the top N corner points with the largest response value as the final corner point set.
4) Maintain the scale invariance by building an image pyramid and testing the corner points on each level of the pyramid.
5) Keep the rotation invariance using the intensity centroid. In each small image block, define the moment of the image block, as shown in Equation (1).

$$m_{pq} = \sum_{x,y \in r} x^p y^q I(x, y) \tag{1}$$

In addition, $I(x, y)$ is the image grayscale of the image block, $p, q = \{0, 1\}$ and the centroid C of the moment is

$$C = (C_x, C_y) = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right). \tag{2}$$

The main direction of feature points is defined as Equation (3):

$$\theta = \arctan \left(\frac{m_{01}}{m_{00}} / \frac{m_{10}}{m_{00}} \right) = \arctan (m_{01}/m_{10}) \tag{3}$$

2) BRIEF DESCRIPTOR CALCULATION

The standard ORB algorithm uses BRIEF to generate feature descriptors, which describe image areas in the binary mode.

This feature can dramatically reduce the contrast among pixels. τ is defined as the test criteria of an $S \times S$ size image area P.

$$\tau(p; x, y) = \begin{cases} 1, p(x) < p(y) \\ 0, p(x) \geq p(y), \end{cases} \tag{4}$$

where $p(x)$ and $p(y)$ are the grayscale at pixels x and y on area P, respectively. When n pairs of test points are chosen, the n-dimensional binary bit string descriptors can be generated with the binary test criteria τ using Equation (5).

$$f_n(p) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; x_i, y_i) \tag{5}$$

The selection of n requires a comprehensive comparison of computational speed, recognition rate, and storage efficiency. Generally, 64,128,256 is selected.

The abovementioned generated feature descriptors lack directivity and require the feature point centroid direction calculated in Equation (3) as the main direction of the BRIEF. For any n sets of binary criteria at x and y, a $2 \times n$ matrix is generated as follows.

$$M = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{bmatrix} \tag{6}$$

The rotation matrix R_θ corresponds to the centroid direction of the feature points, and a directed form M_θ of M is set as $M_\theta = R_\theta M$. The BRIEF with rotation invariant is expressed in Equation (7).

$$g_n(p, \theta) = f_n(p) \mid (x_i, y_i) \in M_\theta \tag{7}$$

3) INITIAL MATCHING

The similarity between the feature points in the ORB algorithm is reflected by the size of the Hamming distance [32]; that is, a small distance indicates high similarity. Furthermore, the bidirectional fast library for approximate nearest neighbors (FLANN) method [33] can be adopted to obtain the initial matches. The RANSAC algorithm [34] is used to refine the obtained matches. The final refined matches can be utilized to acquire the transformation matrixes F (fundamental matrix) and H (homography matrix) between the image pair.

B. TRANSFORMATION MATRIXES EVALUATION

Both the fundamental matrix F and homography matrix H can describe the projection transformation between the image pair. The former is defined by the epipolar constraint that can generate a clear spatial relationship between two matched points, while the latter is more suitable for describing the mapping relationship between two planes. Therefore, the pixel positional relationship between one image pair can be represented by the suitable F matrix and H matrix in accordance with the stereoscopic degree of the scene in the image.

The fundamental matrix F reflects the intrinsic projective geometry of the image pairs and depends only on the camera's internal parameters K and external parameters R and t.

$$F = K^{-T} t^{\wedge} R K^{-1} \tag{8}$$

The projective geometry between a pair of image feature points can be expressed as Equations (9), (10), and (11).

$$p_1 = KP, p_2 = K(RP + t), \tag{9}$$

$$p_2^T K^{-T} t^{\wedge} R K^{-1} p_1 = 0, \tag{10}$$

$$p_2^T F p_1 = 0, \tag{11}$$

where $p_1 (x_1, y_1)$ is the coordinates of a point on the input image, $p_2 (x_2, y_2)$ is the coordinates of its corresponding point on the reference image, and $P = [X, Y, Z]^T$ is the spatial

position of the same point in the input frame coordinate system.

The homography matrix H is typically used to describe the transformation relationship between image pairs from points on a common plane. The relationship between a pair of the corresponding image feature points can be expressed as Equations (12) and (13).

$$p_2 = Hp_1, \tag{12}$$

$$\begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix}, \tag{13}$$

where h_{ij} represents the parameters of the homography matrix H from the reference image to the input image.

To evaluate the calculated fundamental and homography matrices, the evaluation function is proposed in the ORB-slam2 algorithm [29], as expressed in Equations (14) and (15); S_F and S_H correspond to the evaluation scores of the fundamental and homography matrices, and S_M is the unified representation of S_F and S_H.

$$S_M = \sum_i \left(\rho_M\left(d_{cr}^2\left(x_c^i, x_r^i, M\right)\right) + \rho_M\left(d_{rc}^2\left(x_c^i, x_r^i, M\right)\right)\right), \tag{14}$$

$$\rho_M\left(d^2\right) = \begin{cases} \tau - d^2 & d^2 < T_M \\ 0 & d^2 \geq T_M, \end{cases} \tag{15}$$

where d_{cr}^2 and d_{rc}^2 represent the symmetric conversion errors, respectively, d_{cr}^2 is the error from the current image to the reference image, and d_{rc}^2 is the error from the reference image to the current image. T_M is the outlier rejection threshold based on the χ^2 test at 95%, assuming a standard deviation of 1pixel in the measurement error [29].

$$T_H = 5.99, \quad T_F = 3.84$$
$$\tau = T_H \tag{16}$$

When the scene has few planars and a large parallax, the fundamental matrix F is used to calculate the correlation coefficient of the feature point pairs in the two images. Conversely, when the scene contains various planes (or approximate planes) and the parallax is small, the homography matrix H is applied for the calculation. This study uses the evaluation mechanism proposed in ORB-slam2 [29] as Equation (17) to select the matrix model.

$$R_H = \frac{S_H}{S_H + S_F} \tag{17}$$

C. GEOMETRIC CORRESPONDENCE MATCHING

Generally, the standard ORB algorithm can only successfully match most feature points in image matching. The similarity of the descriptors may cause to mismatch results, and some high-quality feature points may fail to match. Therefore, to guarantee the quality of the matching results, the geometric correspondence matching strategy is adopted in this contribution to eliminate wrong matching points and identify as many correct matching points as possible.

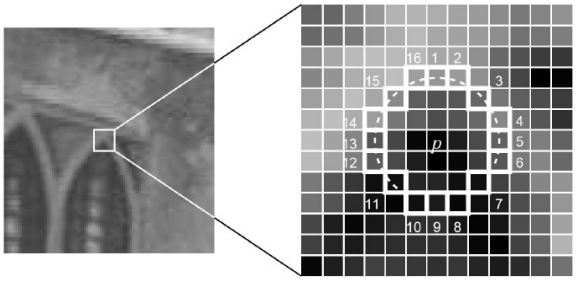


FIGURE 2. FAST feature point [10].

The initial match only uses the local information around the feature points and ignores the overall information in the image. In the present study, the projection transformation relationship of two images is firstly calculated by initial matching, and then the modified NCC method is applied for further feature matching. Traditional rectangular window-based NCC methods are not fixed for rotation and scale changes, but the visual image of the mobile platform frequently has rotation and scale changes. Therefore, we modify the traditional NCC method to improve its robustness in the presence of such a geometric distortion by distorting the correlation window. Specifically, we open a rectangular window in the reference image and project it onto the input image using a transformation matrix obtained in the initial match. This result may be an irregular quadrilateral. Moreover, we use bilinear interpolation to resample the irregular quadrilateral to the same size as the rectangular window. Finally, the correlation coefficient of the rectangular window is calculated in accordance with the standard NCC. The calculation steps are presented as follows:

1) Calculate the position at the input image of the initial feature points that were extracted from the reference image through the fundamental matrix F or the homography matrix H using Equations (11) or (13) in accordance with the matrix at that situation. The selection of the transformation matrix depends on the result of the Equation (17), If $R_H > 0.45$, the homography matrix H is adopted to acquire the correlation coefficient of the feature points. If $R_H \leq 0.45$, we choose the fundamental matrix F. As is shown in Figure 3, a point P on the reference image can be found to have a corresponding point P' on the input image. Due to the deviation in the homography matrix H and the fundamental matrix F, the point closest to P' may not be the corresponding point to P. Thus, the feature points within the n pixels of the retention distance with P' are chosen. If feature points, such as $\{Q_1, Q_2, \cdots, Q_n\}$, are presented and $\{Q_1, Q_2, \cdots, Q_n\}$ is calculated using Equation (18), as shown at the bottom of the next page, and the results are sorted in ascending order. If the correlation coefficient

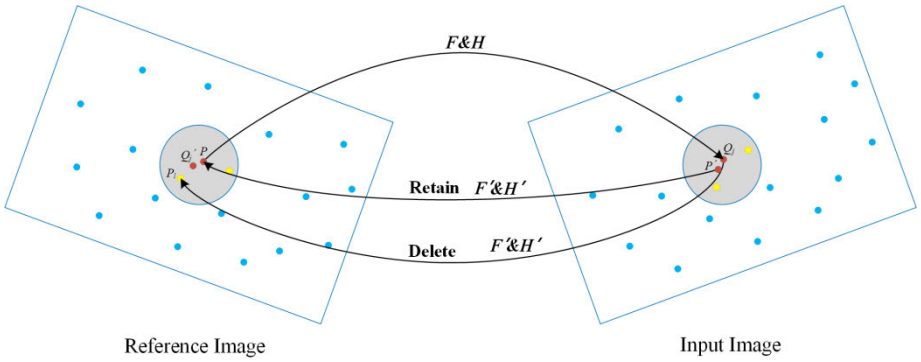


FIGURE 3. Geometric correspondence matching. P is a seed point, P' is its predicted point, Q_j is the matched point, and the other points marked by a yellow circle are those points whose distance to P are within n pixels.

between Q_j and P is the largest one and is greater than r, then the inverse matrix H' or F' of the homography matrix H or the fundamental matrix F are subsequently calculated. Next, the value of the corresponding point Q'_j of Q_j on the reference image using Equation (13) or (11) are calculated. If the distance between P and Q'_j is within n pixels on the reference image and P is the largest correlation coefficient point and is greater than r, then P and Q_j are preliminarily determined as a pair of feature points. In Equation (18), w and h are the width and the height of the matching window, respectively; $g_{i,j}$ is the gray value of the point (i, j) on the reference image; $p(c, r)$ is the correlation coefficient of the reference and input images.

2) Error match elimination is conducted by recalculating the fundamental matrix F_1 or the homography matrix H_1 through the combination of the matching points obtained in the previous step and in the initial matching. Subsequently, the root mean square error (RMSE) is applied to eliminate the wrong matches. If the RMSE is larger than r_2 pixels, then we delete the largest error point by point until the requirements are satisfied. Finally, the mean square errors σ_x and σ_y in the horizontal and vertical directions are calculated, correspondingly. If the horizontal error of the point is greater than E_x or the vertical error is greater than E_y, then the point is removed.
3) With the retained F_1 or H_1 matches as the input, the previous steps are repeated to increase the number of correct matches. In general, after several iterations, the number of matching results remains essentially the same. Given the time requirement, we double the maximum number of repetitions.

III. EXPERIMENTS AND ANALYSIS

In this study, the standard ORB algorithm and the initial matching algorithm in ORB-slam2 are compared with our proposed GC-ORB algorithm. The specific process of the ORB-slam2 initial matching algorithm is presented as follows: This method rasterizes the entire image to improve the match efficiency and modifies the minimum Hamming matching distance to 0.7–0.9 times of the sub-optimal matching distance. Moreover, this method uses the histogram statistics to describe the change in the main direction angle in the rotation of the descriptors and selects only the top three feature points of statistical ranking to improve the matching accuracy of the algorithm.

The testing environment for this experiment is conducted in the personal computer involving Intel Core i7 CPU 2.75 GHz, and 16 GB memory. All three algorithms were programmed using C++ with the OpenCV3.4.3 in the Ubuntu16.04 environment.

A. THE DESCRIPTION OF EXPERIMENTAL DATASETS

To evaluate the performance of the algorithm, three sets of RGB-D images from two public datasets, namely, ICL-NUIM and TUM are used. The RGB-D datasets contain more accurate depth information than the triangulation algorithm, with a pose supported by the datasets that can help to verify the matching accuracy of the three algorithms.

The TUM datasets capture color and depth images through the Microsoft Kinect sensor and also contain the pose. The

$$p(c,r)=\frac{\sum_{i=1}^{w}\sum_{j=1}^{h}g_{i,j}\cdot g'_{i+r,j+c}-\frac{1}{w\cdot h}\left(\sum_{i=1}^{w}\sum_{j=1}^{h}g_{i,j}\right)\left(\sum_{i=1}^{w}\sum_{j=1}^{h}g'_{i+r,j+c}\right)}{\sqrt{\left[\sum_{i=1}^{w}\sum_{j=1}^{h}g_{i,j}^{2}-\frac{1}{w\cdot h}\left(\sum_{i=1}^{w}\sum_{j=1}^{h}g_{i,j}\right)^{2}\right]\left[\sum_{i=1}^{w}\sum_{j=1}^{h}g_{i+r,j+c}^{\prime 2}-\frac{1}{w\cdot h}\left(\sum_{i=1}^{w}\sum_{j=1}^{h}g'_{i+r,j+c}\right)^{2}\right]}} \tag{18}$$

TABLE 1. Experimental image pairs.

No.	Figure	Data Source	Resolution	Noise	Motion Blur
1	4a 4b	TUM	640×480	√	
2	4c 4d	TUM	640×480	√	√
3	4e 4f	ICL-NUIM	640×480		

collection rate of the data is 30 Hz, and the sensor resolution is 640 × 840 pixels. The ground truth is obtained through the eight-lens high-precision motion capture system.

The ICL-NUIM datasets are obtained from Imperial College London. This dataset is designed to provide accurate RGBD data for visual odometers and SLAM algorithms, thereby providing different scenes with ground truth collected using Microsoft Kinect sensors. The collection rate of the data is 30 Hz, and the sensor resolution is 640 × 840.

The two datasets in the TUM dataset contain an empty shaded scene and a desktop with a complex texture scene, thus providing complex and monotonous scene textures. The ICL-NUIM dataset offers noise-free scenes. We use three sets of image pairs from different scenarios (Table 1) to evaluate the performance of the algorithm in different scenarios.

The groundtruth of the datasets exists between adjacent frames or two frames with the specified interval. It contains pose and depth information that are the essential basis for assessing the quality of experimental results. Table 2 contains the pose relationship between the two images of the three image pairs. All the pose information is represented by the quaternion. The depth images, due to their weak discrimination, are not shown in this paper.

B. THE PARAMETER SETTINGS

1) PARAMETERS IN THE INITIAL ORB FEATURE EXTRACTION AND MATCH

The number of initial extracting feature points is set as a maximum of 500 points. The scale between the pyramid images is 1.2f (float). The number of layers in the Gaussian pyramid is 8, and the index value of the first level is 0. The edge threshold is set as 31 pixels, and the number of point pairs for producing BRIEF descriptors is set as 2. The feature point field size of the BRIEF descriptor is set as 31.

2) PARAMETERS IN THE FUNDAMENTAL AND HOMOGRAPHY MATRIXES

The symmetric transformation error T_F of the fundamental matrix is 3.82, and the symmetric transformation error T_H of the homography matrix is 5.99. The conversion threshold R_H between the fundamental and homography matrices is set to 0.45. When $R_H > 0.45$, the homography matrix is chosen for the geometric correspondence matching. Conversely, when $R_H \leq 0.45$, the fundamental matrix is selected for such a matching algorithm.

3) PARAMETERS IN THE GEOMETRIC CORRESPONDING MATCHING

In the preliminary experiment, the correct values of n and r are identified by first assigning different values to n from 0.5 to 3.0 in increments of 0.5 and similarly from 0.6 to 0.9. The value of r is evaluated in increments of 0.1. The results of several experiments show that the cost function of the final transformation matrixes can constantly be maintained within 1.5 pixels when $n = 1$ and $r = 0.8$. Thus, the matching points can be considered correct matches. In the next step, r_1 is set as 1 to effectively eliminates false matches. E_x and E_y are set as $3\sigma x$ and $3\sigma y$. If the matching horizontal or vertical error is greater than $3\sigma x$ and $3\sigma y$, then this result can be regarded as a mismatch. This method may eliminate some correct matches and ensure the high precision of the matching points.

C. THE EVALUATION CRITERIA AND IMPLEMENTATION DETAILS

The matching quality of the algorithm is evaluated using the following factors: the number of correct matches, the correct matching rate, the matching accuracy, and the time consumption.

1) THE NUMBER OF CORRECT MATCHES

For all the matched feature point pairs, if any of them has a residual error higher than 1.7 pixels, then this condition is considered as an error match; otherwise, a correct match is obtained.

2) THE CORRECT MATCHING RATE

The number of the correct matching point pairs divided by the number of all matching feature point pairs is the correct matching rate.

3) THE MATCHING ACCURACY

The evaluation of matching accuracy generally requires ground truth. For image matching, the result of the manual measurement is commonly used as a reference value. However, the accuracy of counting such numerous feature points manually and maintaining the quality of the results depend on the accuracy of the individual surveyor. This facet is subjective and varied from person to person. Therefore, a more practical and objective approach is required. In this study, a matching accuracy evaluation method based on RMSE is adopted. We use feature matching to match the image pairs, and each point has a residual from which the RMSE is calculated. Theoretically, the RMSE consists of the following three parts: the image matching error, the image modeling error, and the image internal geometric distortion. The image matching error is the major component of the RMSE error. The two other errors are secondary components and can be ignored for images with improved geometric

TABLE 2. Pose information between image pairs.

	tx/m	ty/m	tz/m	qx	qy	qz	qw
4a 4b	-0.00013625	-0.00000146	0.00003787	0.000545849	0.00152376	-0.00000656	0.999999
4c 4d	-0.00681431	-0.00215337	0.00747452	-0.000071321	0.00513401	-0.00143419	0.999986
4e 4f	-0.00121954	-0.00143325	-0.00089346	0.000101122	0.00095517	-0.00037221	0.999999

TABLE 3. Experimental results of standard oriented fast and rotated BRIEF (ORB), ORB-slam2 initial matching algorithm and the proposed method, geometric correspondence - oriented FAST and rotated BRIEF (GC-ORB).

No.	Algorithm	Extracted Features		Correct Matches	Correct Matching Rate	RMSE/ pixel	Time/s
		Reference Image	Input Image				
1	ORB	484	490	143	82.08%	2.6782	0.012
	ORB-slam2	505	506	36	97.30%	1.6545	0.021
	GC-ORB	**484**	**490**	**193**	**99.48%**	**1.4971**	**0.046**
2	ORB	500	500	153	52.34%	2.3510	0.016
	ORB-slam2	508	507	25	54.35%	2.9512	0.021
	GC-ORB	**500**	**500**	**212**	**69.05%**	**2.0955**	**0.045**
3	ORB	500	500	255	84.27%	2.5283	0.013
	ORB-slam2	505	504	46	83.64%	1.9713	0.018
	GC-ORB	**500**	**500**	**324**	**97.88%**	**1.3191**	**0.044**

quality. Therefore, the RMSE can partially objectively reflect the image matching accuracy, although the actual positional accuracy of the matching points can be better than the RMSE.

4) THE TIME CONSUMPTION

The algorithm initially extracts the feature points and then provides the descriptors for the feature points; finally, the algorithm outputs the matching point pairs. The time consumed from all the steps is the algorithm time consumption. In general, when the time is less than 0.05s, the algorithm can satisfy the real-time requirements. The time consumption of the algorithm also depends on the hardware conditions of the running platform, thus possibly reflecting the efficiency of the algorithm slightly.

5) THE COMPARATIVE RESULTS AND ANALYSIS

Table 3 lists the number of feature points, the correct matching points, the correct matching rate, the matching accuracy RMSE and the matching time of the standard ORB algorithm, the ORB-slam2 initial matching algorithm, and the GC-ORB algorithm, respectively.

The experimental results listed in Table 3 indicate that the ORB-slam2 initial matching algorithm matches only a few feature point pairs after extracting a substantially similar number of feature points. This result is due to the algorithm performs histogram statistics in accordance with the direction of the sub-rotation, and only the top three main direction points are retained when filtering feature points. This screening mechanism eliminates many feature point pairs. If the other matching methods are not used, the attitude information between the image pair may be restored with numerous errors. Conversely, the GC-ORB algorithm proposed in this study eliminates mismatched pairs through the geometric correspondence between the image and the surrounding matching points. The standard ORB is characterized by fast speed, with an advantage to ensure satisfactory real-time performance for mobile devices in image acquisition and the feature point method. The GC-ORB algorithm sacrifices part of the speed and improves the matching accuracy to maintain the real-time performance of the algorithm. In addition, the proposed algorithm maintains or slightly increases the number of accurate matching point pairs. Based on the same or even higher matching accuracy, the number of matching point pairs is several times higher in the GC-ORB algorithm than that in the initial matching algorithm in ORB-slam2.

To some extent, the RMSE denotes that the standard ORB algorithm has some mismatch results. Although ORB-slam2 removes some mismatched points in accordance with the main direction of the descriptor rotation, the FAST feature points are accompanied by many edge features. In various scenarios, the large-scale edge features of objects have similar situations, while the descriptors cannot distinguish the differences, thereby causing multiple feature points extracted on the same edge of the object to match incorrectly. Given that the ORB-slam2 initial algorithm is strictly limited by the threshold, only few extracted feature point pairs are few. Thus, the large errors in the individual point pairs may result in a significant RMSE value. The experimental results show that the RMSE of the GC-ORB matching can constantly maintain high precision.

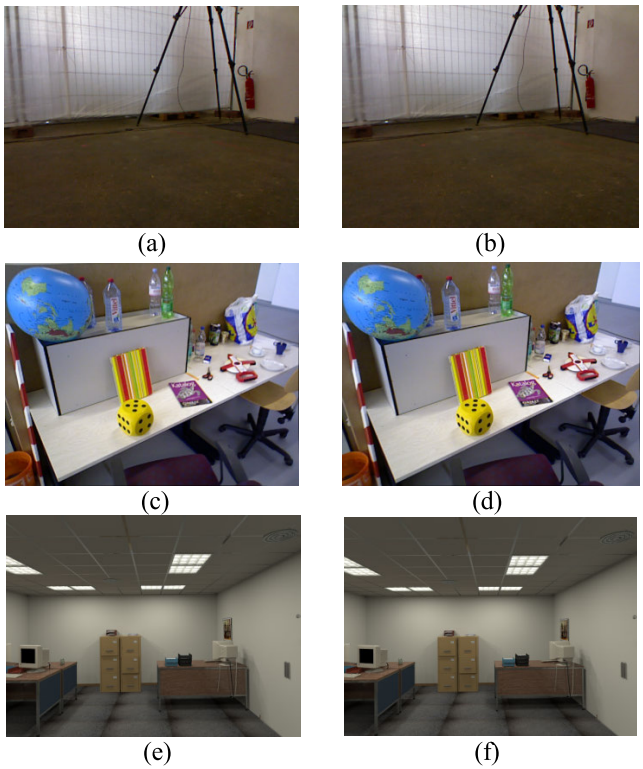


FIGURE 4. Experimental visual image pairs. (a-b): empty shaded scene from TUM; (c-d): complex texture with motion blur from TUM; (e-f): regular scene without noise from ICL-NUIM.

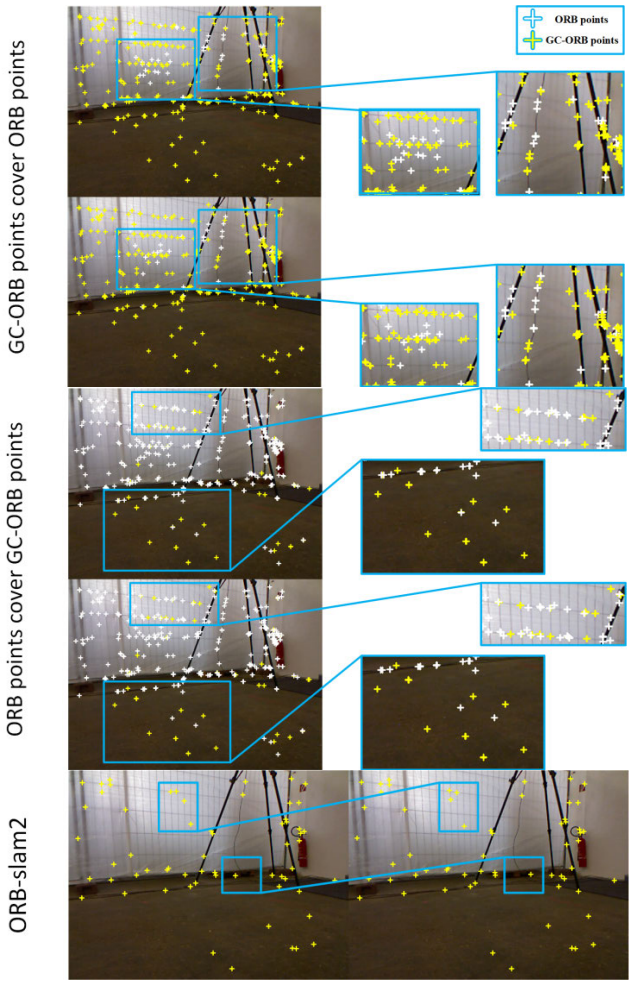


FIGURE 5. Matching results for the first experimental image pair: Left and Right images are both from TUM datasets.

In terms of the time consumption, the GC-ORB is two to three times slower than the standard ORB and ORB-slam2 initial matching algorithms but the time consumption can be maintained under 0.05s. The GC-ORB algorithm exerts time in the geometric correspondence matching, a factor which is chiefly affected by the number of matching points obtained in the initial matching. With the adjustment in the threshold number of feature point extraction, the time consumption of the GC-ORB algorithm increases rapidly (later demonstrated in Figure 11). However, the time consumption of the algorithm also depends on the performance of hardware devices and the optimization of the algorithm. With the development of hardware platforms in the future, the time consumption of the GC-ORB algorithm can be decreased on improved platforms.

In this study, when the matching results of the GC-ORB algorithm in each group of images are displayed, the GC-ORB matching points are shown on the same pair of images to cover the ORB matching point pairs, and the ORB matching points are also shown on the same images to cover the GC-ORB matching point pairs. The former can reflect the points (white points) eliminated by the GC-ORB from the final result, and the latter can reflect the additional points (yellow points) added by the GC-ORB at the subsequent algorithm steps.

Figure 5 demonstrates the matching results of the first pair of TUM data with large shadow areas collected at an open field. The texture structure in the image is relatively simple. Empty patch features and a grid consisting of numerous line features are included in the image. In Figure 4, the ORB algorithm has good distribution in the shadow texture part because the shadow portion of the image has favorable line feature textures. The selection of FAST feature points only compares the brightness difference among pixels. Thus, when the algorithm extracts the feature points, the FAST corner points gather many line features. For the patch area with uniform brightness and color, only a few or even no feature points are obtained. However, the FAST feature points clustered on the line features are overly similar in brightness and grayscale. This outcome is typical when using the ORB algorithm for feature extraction and matching, especially when many or long line features are found in the image. The ORB-slam2 initial matching algorithm removes most mismatched points through the rotation direction of the feature point descriptor but still cannot guarantee the eradication of all mismatched points, as exhibited in the matching results in Figures 5, 6, and 7. Through the geometric relationship between image pairs, GC-ORB removes nearly

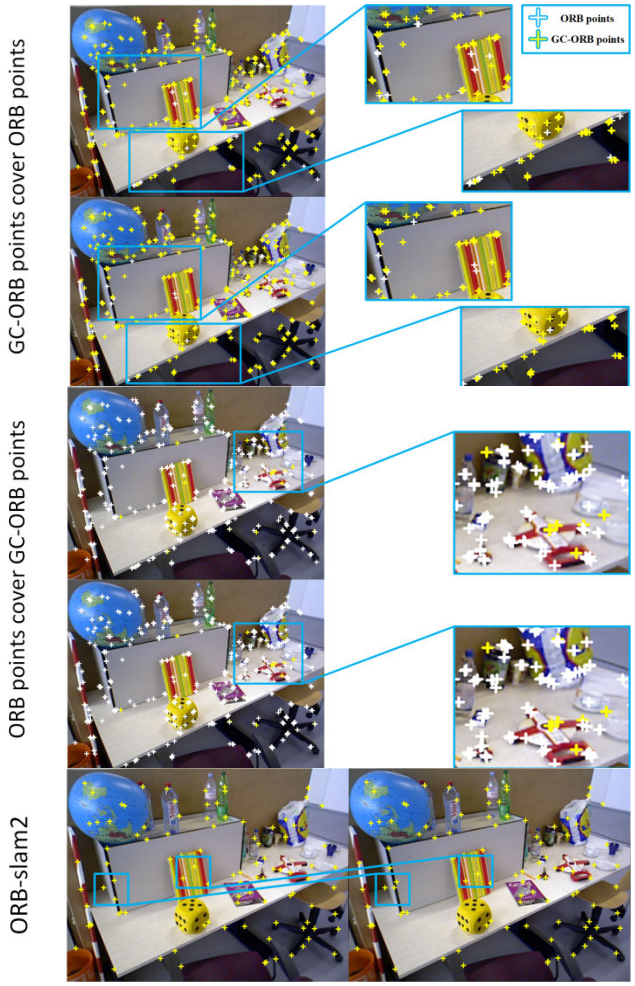


FIGURE 6. Matching results for the second experimental image pair: Left and Right images are both from TUM datasets.

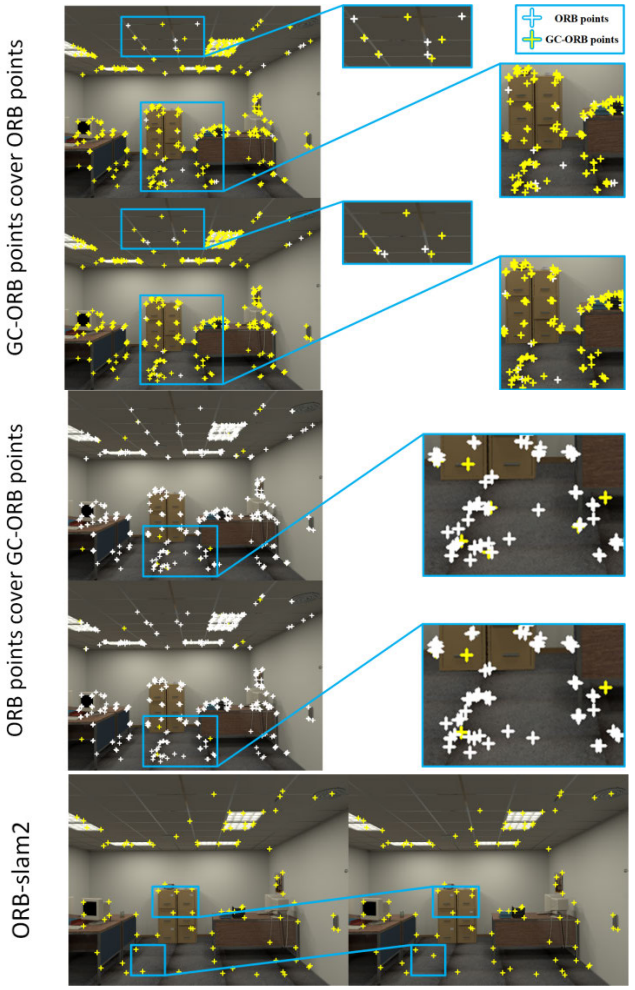


FIGURE 7. Matching results for the third experimental image pair: Left and Right images are both form ICL-NUIM datasets.

all mismatched points from the matching pairs of points extracted by the standard ORB algorithm and identifies new matching points. The results show that GC-ORB achieves a more accurate match than ORB based on maintaining the number of matching points and adds numerous accurate matches based on maintaining or exceeding the accuracy of the ORB-slam2 initial matching. Thus, the GC-ORB can be applied to an open area where the texture is monotonous.

Figure 6 displays the matching results of a second pair of TUM data images with large complex objects collected in a small office space. Many kinds of objects are found in the image, and the texture structure is complex. Given the close distance between the lens and the object, a dynamic blur occurs in the collected image. This phenomenon results in a large displacement deviation between the point pairs of the three matched methods. Moreover, the RMSE of the three algorithms exceeds two pixels. The matching accuracy in this experiment is evaluated on the basis of the displacement deviation between the matching and the true points. Therefore, under our experimental conditions and evaluation criteria, the matching accuracy of the three algorithms in the second pair of images has decreased. In a scene with complex texture structure, ORB has a good distribution but is influenced by the presence of line features in the scene, and the phenomena of feature point gathering and mismatching still exist. The result of the ORB-salm2 initial matching algorithm is also affected by the same factor and has some mismatched point pairs. GC-ORB eliminates mismatched points based on the ORB, improves the correct matching rate and accuracy, and reduces the RMSE of the overall point cloud. Thus, the GC-ORB method is applicable to complex scene regions and can be used in the case of dynamic blurring of acquired images. Relatively improved results are obtained through this technique.

Figure 7 illustrates the matching results of the third pair of ICL-NUIM data images collected in a noise-free indoor field. The results include shadows, desktops, ceilings, and other textures. The standard ORB and ORB-slam2 initial matching algorithms still have mismatches on the metal edge of the ceiling and on the shadow of the carpet. Conversely, the GC-ORB eliminates these mismatches. In general, GC-ORB has a

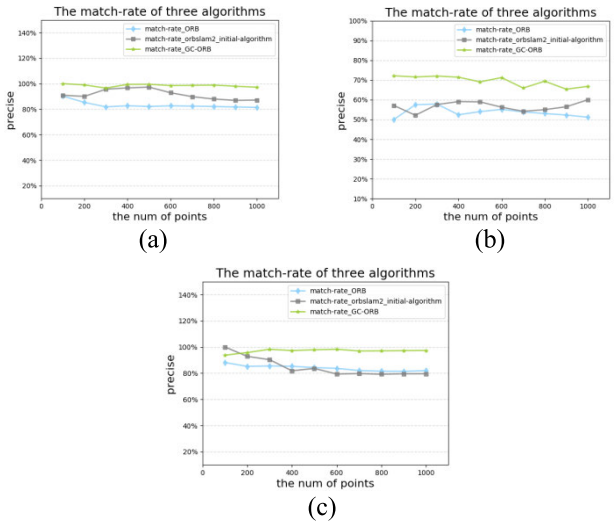


FIGURE 8. Correct matching rate of three algorithms in three test image pairs form 100 feature points to 1000 feature points: (a) empty shaded scene from TUM datasets; (b) complex texture with motion blur from TUM datasets; (c) regular scene without noise from ICL-NUIM datasets.

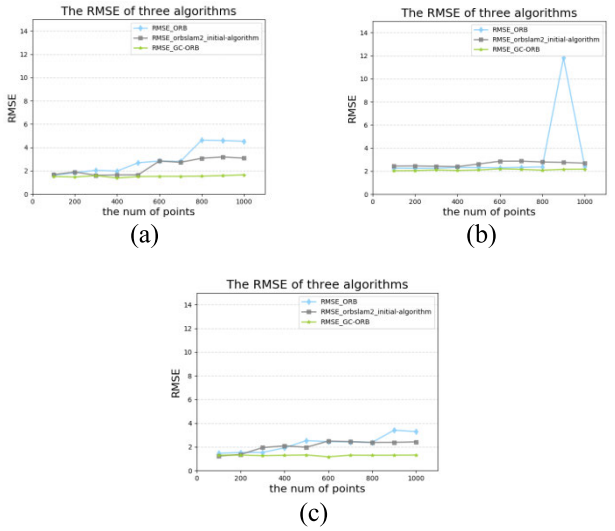


FIGURE 9. RMSE of three algorithms in three test image pairs form 100 feature points to 1000 feature points: (a) empty shaded scene from TUM datasets; (b) complex texture with motion blur from TUM datasets; (c) regular scene without noise from ICL-NUIM datasets.

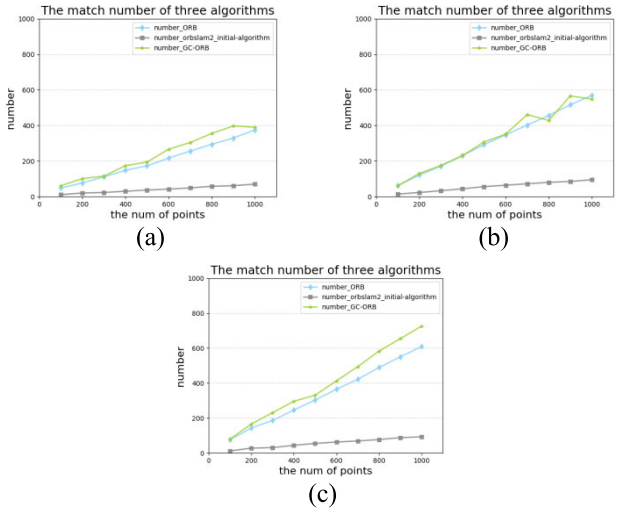


FIGURE 10. The match number of three algorithms in three test image pairs form 100 feature points to 1000 feature points: (a) empty shaded scene from TUM datasets; (b) complex texture with motion blur from TUM datasets; (c) regular scene without noise from ICL-NUIM datasets.

higher matching accuracy than ORB and ORB-slam2 initial matching algorithms, thereby indicating that the method can be applied to various indoor scenes.

Based on existing experiments, the threshold number of feature points is set to 100 increments from 100 to 1000. The correct matching rate, RMSE, matching number, and time consumption are compared to determine the difference of the number of feature points extracted through the algorithms.

Figure 8 gives the variation in the correct matching rate of the three algorithms for the three pairs of data with different numbers of feature points as the initial state. The comparison of the three algorithms can basically maintain the stability of the correct rate when the feature points are increased, and the correct rate of the GC-ORB algorithm outperforms those of the two other algorithms.

Figure 9 demonstrates the variation in the RMSE for the three algorithms with the increase in the number of feature points. Among these three sets of data, the GC-ORB algorithm maintains the highest precision. In the second set of data exhibited in Figure 9 (b), the texture structure is complex, and the line features present numerous cases with small sizes. The initial matching points obtained through the ORB algorithm are evenly distributed, and the number of extreme mismatches is relatively small. Thus, with the increase of the number of feature points, these three algorithms can ensure smoothness. Under the condition of 900 feature points, the ORB algorithm may match the wrong point pairs with large differences in individual positions, thereby resulting in a sharp increase in the RMSE. The general trend shown by the three datasets indicates that the RMSE of the ORB and ORB-slam2 initial matching algorithms will gradually increase with the number of initial feature points. Since the GC-ORB algorithm removes the point pairs with large position distances in the matching step, the change in the number of feature points slightly affects its RMSE of the matching results.

Figure 10 displays the variation in the number of matching results when the number of feature points in the three algorithms increases. The three sets of image line graphs show that the matching result obtained through the GC-ORB algorithm can maintain the highest number among the three algorithms with the increase in the number of feature points. In the subsequent algorithm of slam, the number of feature point matching results may affect the smoothness of the recovery trajectory. High matching results and uniform distributions imply a smooth trajectory result of the mobile platform.

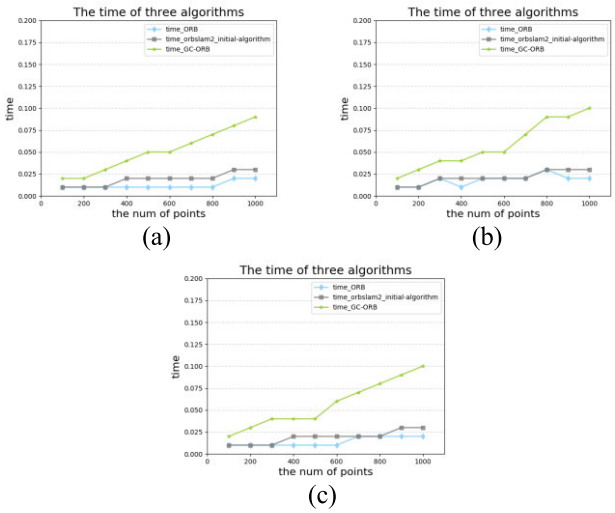


FIGURE 11. Time-consuming of three algorithms in three test image pairs form 100 feature points to 1000 feature points: (a) empty shaded scene from TUM datasets; (b) complex texture with motion blur from TUM datasets; (c) regular scene without noise from ICL-NUIM datasets.

Figure 11 illustrates the time consumption changes of these three algorithms with the increase in the number of feature points. The three sets of data line graphs indicate that the GC-ORB algorithm can be kept within 0.05 s while maintaining the 500 feature points, a situation which can satisfy the requirements of real-time performance. When the number of feature points exceeds 600, the time consumption of the algorithm is significantly increased because the algorithm must count the displacement errors of all pairs of initial matching feature points in the geometric matching step. A high number of pairs suggests that the algorithm is time-consuming. However, the GC-ORB algorithm can essentially keep the time within 0.1s. Thus, the GC-ORB algorithm can be considered for the extraction and matching of keyframe feature points.

IV. DISCUSSION

This study presents GC-ORB, which is an effective mobile platform image matching method, and evaluates its performance using three sets of data sources in varied scenes. In accordance with the evaluation index of the matching number, the correct matching rate, the RMSE, and the time consumption, two common methods (i.e., the standard ORB and the ORB-slam2 initial matching algorithms) are compared and analyzed in relation to the GC-ORB method. The experimental results show that, based on real-time requirements, the GC-ORB algorithm has higher correct matching rate and matching accuracy in accordance with the number of similar matches than the original ORB algorithm. Moreover, the GC-ORB algorithm has more matching pairs based on similarity and a higher correct matching rate and matching accuracy than the ORB-slam2 initial matching algorithm.

Although the original ORB algorithm can extract sufficient feature points independently and quickly, the matching results of the ORB constantly aggregate and produce mismatches on the line features with similar textures given the characteristics of FAST corner points. The number of mismatches depends on the number and length of line features in the scene. Mismatching affects the attitude recovery and tracking of the mobile platform in the late stage. The ORB-slam2 initial matching algorithm improves feature point matching efficiency by rasterizing the entire image, modifies the minimum Hamming matching distance to 0.7–0.9 times the sub-optimal matching distance, uses histogram statistics to describe the sub-rotation main direction angle change value, and selects only the top three feature points of the statistical ranking to increase the matching precision of the feature points. However, given the similar line characteristics of the texture, a mismatch still occurs in the small number of high-precision matching point pairs matched by the ORB-slam2 initial matching algorithm. Considering that the extracted matching points are even and rare, the mismatched points in the algorithm significantly affect the correct matching rate and matching accuracy. The pose recovery of the mobile platform will be increasingly affected and may cause the platform to obtain a path that is insufficiently smooth or one that loses tracking.

To obtain sufficient matching and satisfactory distribution, we first use the FAST operator to extract sufficient feature points quickly, thereby laying a good foundation for the subsequent stage of feature matching. In the initial matching, we generate many matching point pairs and the fundamental and the homography matrices between image pairs. Furthermore, we use the evaluation function to filter the matrix that is suitable for the current scene. To maximize the FAST corner points, this study performs the geometric correspondence matching based on the selected transformation matrix using the geometric relationship between image pairs rather than partial image information. This process constrains the corresponding points in a small search area to find increasingly accurate matching points. These matching points are evenly distributed and highly reliable and have robust-to-weak textures or motion-blurred images.

To ensure high accuracy, the GC-ORB implements multiple constraints throughout the matching process. In the initial matching process, FLANN, Hamming distance with filtering threshold, and RANSAC with a strict value are used to ensure that the initial matching is correct. In the geometric correspondence matching process, we use RMSE in the global, X, and Y directions to eliminate mismatches.

Our experimental image pairs cover different scenes, including open, shaded, complex indoor, close-range dynamic blurred, and normal indoor scenes. The GC-ORB performs well on all scene images. In summary, the GC-ORB is effective and stable for feature matching of visual images acquired on small mobile platforms.

V. CONCLUSION

This study proposes a multi-level matching method for visual images acquired using mobile platforms. The suggested

method includes the following steps: FAST feature point extraction, initial matching, transformation matrices evaluation, and geometric correspondence matching. The main contribution of this study is to propose a matching selection strategy. Geometric correspondence matching uses the geometric transformation information between images to identify matches that satisfy geometric transformations. Comprehensive experiments on visual images with different angles of view, ambiguity, and texture complexity show that the method improves the correct matching rate and accuracy while maintains a real-time performance.

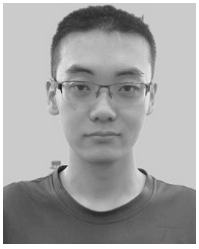
The method proposed in this study can be applied to various visual image applications that require feature point matching. For future work, different visual images will be adopted to evaluate the robustness of the suggested technique, and we will aim to optimize the algorithm content and enhance the efficiency of the algorithm.

REFERENCES

[1] A. Gruen, "Development and status of image matching in photogrammetry," *Photogram. Rec.*, vol. 27, no. 137, pp. 36–57, Mar. 2012.
[2] L. Yan, D. R. Roy, H. Zhang, J. Li, and H. Huang, "An automated approach for sub-pixel registration of landsat-8 operational land imager (OLI) and sentinel-2 multi spectral instrument (MSI) imagery," *J. Remote Sens.*, vol. 8, no. 6, p. 520, Jun. 2016.
[3] M. Gianinetto, "Automatic co-registration of satellite time series," *J. Photogramm. Rec.*, vol. 27, pp. 462–470, Dec. 2012.
[4] S. L. Zhou, X. Z. Wu, and G. Liu, "Integrated navigation method of monocular ORB-SLAM/INS," *J. Chin. Iner. Tech.*, vol. 24, no. 5, pp. 633–637, May 2016.
[5] X. Lin, F. Wang, L. Guo, and W. Zhang, "An automatic key-frame selection method for monocular visual odometry of ground vehicle," *IEEE Access*, vol. 7, pp. 70742–70754, 2019.
[6] J. Feddema and O. R. Mitchell, "Vision-guided servoing with feature-based trajectory generation," *IEEE Trans. Robot. Automat.*, vol. RA-5, no. 5, pp. 691–700, Oct. 1989.
[7] Y. Zhang and Z. Miao, "Object recognition based on ORB and self-adaptive kernel clustering algorithm," in *Proc. IEEE Int. Conf. Signal Process.*, Hangzhou, China, Oct. 2014, pp. 1397–1402.
[8] N. Mahmoud, I. Cirauqui, A. Hostettler, C. Doignon, L. Soler, J. Marescaux, and J. M. M. Montiel, "ORB-SLAM-based endoscope tracking and 3D reconstruction," in *Proc. Int. Workshop Comput. Assist. Robot. Endoscopy*, Cham, Switzerland: Springer, 2016, pp. 72–83.
[9] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571.
[10] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2006, pp. 430–443.
[11] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BRIEF: Binary robust independent elementary features," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Hersonissos, Greece, Sep. 2010, pp. 778–779.
[12] M. Calonder, V. Lepetit, M. Ozuysal, T. Trzcinski, C. Strecha, and P. Fua, "BRIEF: Computing a local binary descriptor very fast," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 7, pp. 1281–1298, Jul. 2012.
[13] Y. Qin, H. Xu, and H. Chen, "Image feature points matching via improved ORB," in *Proc. IEEE Int. Conf. Progr. Informat. Comput. (PIC)*, Shanghai, China, May 2014, pp. 204–205.
[14] Y. Zhu, X. Shen, and H. Chen, "Copy-move forgery detection based on scaled ORB," *Multimedia Tools App.*, vol. 75, no. 6, pp. 3221–3233, Mar. 2015.
[15] X. Wang, J. Zou, and D. Shi, "An improved ORB image feature matching algorithm based on SURF," in *Proc. 3rd Int. Conf. Robotic. Autom. Eng. (ICRAE)*, Guangzhou, China, Nov. 2018, pp. 218–222.
[16] C. Feng, Z. Yang, L. Wang, and Y. Li, "Improved mar natural feature recognition algorithm based on surf and ORB," in *Proc. IEEE Int. Conf. Sensor Netw. Signal Process. (SNSP)*, Xi'an, China, Oct. 2018, pp. 462–468.
[17] R. Wang, W. Zhang, Y. Shi, X. Wang, and W. Cao, "GA-ORB: A new efficient feature extraction algorithm for multispectral images based on geometric algebra," *IEEE Access*, vol. 7, pp. 71235–71244, 2019.
[18] H. Yu, Q. Fu, Z. Yang, L. Tan, W. Sun, and M. Sun, "Robust robot pose estimation for challenging scenes with an RGB-D camera," *IEEE Sensors J.*, vol. 19, no. 6, pp. 2217–2229, Mar. 15, 2019.
[19] W. Fang, Y. Zhang, B. Yu, and S. Liu, "FPGA-based ORB feature extraction for real-time visual SLAM," in *Proc. IEEE Int. Conf. Field Program. Technol. (ICFPT)*, Melbourne, VIC, Australia, Dec. 2017, pp. 275–278.
[20] P. Tran, T. H. Pham, S. K. Lam, M. Wu, and B. A. Jasani, "Stream-based ORB feature extractor with dynamic power optimization," in *Proc. Int. Conf. Field-Program. Technol. (FPT)*, Naha, Okinawa, Japan, Dec. 2018, pp. 94–101.
[21] C. Ma, X. Hu, L. Fu, and G. Zhang, "An improved ORB algorithm based on multi-feature fusion," in *Proc. IEEE 27th Int. Symp. Ind. Electron. (ISIE)*, Cairns, QLD, Australia, Jun. 2018, pp. 729–734.
[22] J. Tang, L. Ericson, J. Folkesson, and P. Jensfelt, "GCNv2: Efficient correspondence prediction for real-time SLAM," 2019, *arXiv:1902.11046*. [Online]. Available: https://arxiv.org/abs/1902.11046
[23] Z. Yu, K. Wang, and R. Li, "The inter-frame feature matching and tracking of binocular vision based on the ORB-PyrLK algorithm," in *Proc. IEEE 36th Chin. Control Conf. (CCC)*, Dalian, China, Jul. 2017, pp. 6922–6927.
[24] G.-X. Xin, X.-T. Zhang, X. Wang, and J. Song, "A RGBD SLAM algorithm combining ORB with PROSAC for indoor mobile robot," in *Proc. IEEE 4th Int. Conf. Comput. Sci. Netw. Technol. (ICCSNT)*, Harbin, China, Dec. 2015, pp. 71–74.
[25] C. Shu and X.-Z. Xiao, "ORB-oriented mismatching feature points elimination," in *Proc. IEEE Int. Conf. Progr. Informat. Comput. (PIC)*, Suzhou, China, Dec. 2018, pp. 246–249.
[26] L. Oi, W. Liu, and D. Liu, "ORB-based fast anti-viewing image feature matching algorithm," in *Proc. IEEE Chin. Autom. Congr. (CAC)*, Xi'an, China, Dec. 2018, pp. 2402–2406.
[27] Y. Wang and S. Ge, "An improved ORB image matching algorithm based on compressed sensing," in *Proc. 8th Int. Conf. Comput. Eng. Netw. (CENet)*, vol. 905. Cham, Switzerland: Springer, 2018, pp. 351–359.
[28] J. Yan, Z. Wang, and S. Wang, "Real-time tracking of deformable objects based on MOK algorithm," *J. Syst. Eng. Electron.*, vol. 27, no. 2, pp. 477–483, Apr. 2016.
[29] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós, "ORB-SLAM: A versatile and accurate monocular SLAM system," *IEEE Trans. Robot.*, vol. 31, no. 5, pp. 1147–1163, Oct. 2015.
[30] X. Geng, "Research on photogrammetric processing for Mars topographic mapping," *Acta Geodaetica Et Cartographica Sinica*, vol. 44, no. 8, p. 944, Aug. 2015.
[31] Y. Liu, F. Mo, and P. Tao, "Matching multi-source optical satellite imagery exploiting a multi-stage approach," *Remote Sens.*, vol. 9, p. 1249, Nov. 2017.
[32] R. W. Hamming, "Error detecting and error correcting codes," *Bell Syst. Tech. J.*, vol. 29, no. 2, pp. 147–160, Apr. 1950.
[33] M. Muja and D. G. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," in *Proc. Int. Conf. Comput. Vis. Theory Appl. VISAPP*, vol. 1, no. 2, pp. 331–340, Feb. 2009.
[34] M. A. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

YU CHEN received the master's degree from Information Engineering University, Zhengzhou, China, in 2017. He is currently pursuing the Ph.D. degree with the School of Geodesy and Geomatics, Wuhan University, Wuhan, China. His research interests include multisensor integrated navigation and UAV autonomous route planning, and VIO/VSLAM.

LI YAN received the B.S., M.S., and Ph.D. degrees in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1989, 1992, and 1999, respectively. He is currently a Luojia Distinguished Professor with the School of Geodesy and Geomatics, Wuhan University. His research interests include photogrammetry, remote sensing, and precise image measurement.



BO XU received the B.S. degree from the China University of Geosciences Beijing, Beijing, China, in 2018. He is currently pursuing the master's degree with the School of Geodesy and Geomatics, Wuhan University, Wuhan, China. His research interests include visual-inertial odometery and integrated navigation.

YUXUAN LIU received the master's degree from Information Engineering University, Zhengzhou, China, in 2016. He is currently pursuing the Ph.D. degree with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China. His research interests include image matching, and 3D reconstruction from images and point clouds.