IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 30, 2018, accepted September 18, 2018, date of publication September 26, 2018, date of current version October 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2871729

UGC: Real-Time, Ultra-Robust Feature Correspondence via Unilateral Grid-Based Clustering

ZHAOHUI ZHENG[1,2], YONG MA[1], HONG ZHENG[1], JIANPING JU[1,3], AND MINGYU LIN[3]

[1]School of Electronic Information, Wuhan University, Wuhan 430079, China
[2]Department of Public Courses, Wuhan Railway Vocational College of Technology, Hubei 430205, China
[3]School of Mechanical & Electrical and Information Engineering, Hubei Business College, Hubei 430000, China

Corresponding author: Hong Zheng (zh@whu.edu.cn)

This work was supported by the Shenzhen Basic Science & Technology Foundation of China under Grant JCYJ20150422150029095.

ABSTRACT Quickly establishing reliable correspondence between two feature sets is a challenging task for feature matching. However, the key to successful feature matching is not only matching robustness but also the precision and real-time performance. It is difficult to achieve both efficiency and efficacy using the current algorithms. In this paper, we propose unilateral grid-based clustering (UGC), which creates a unilateral grid of an image's features and meanshift clustering constraints of the other image correspondence features. UGC removes a large number of mismatches using clustering center statistical analysis of the match feature points in a grid region. For low texture, blur and wide-baselines feature matching of images, UGC provides a real-time, ultra-robust correspondence system. Extensive experiments on image data sets demonstrate the higher precision and real-time performance of UGC, which outperforms current state-of-the-art methods, including conditions such as low contrast and high exposure.

INDEX TERMS Feature matching, unilateral grid-based clustering, real-time, ultra-robust, correspondence, mismatch.

I. INTRODUCTION

Image registration is a fundamental and challenging computer vision processing problem. The primary purpose of image registration is matching and overlaying two or more images acquired at different times using different sensors (imaging equipment) or under different conditions (weather, illumination, camera position, and angle). It has been widely used in many fields, including remote sensing data analysis, computer vision, and image processing [1], [2]. Feature matching is an important branch of image registration, and a number of feature registration methods have been produced [3], [4]. In general, these methods are oriented to a certain range of applications and have their own characteristics; thus, it is difficult to integrate robustness, accuracy, and real-time performance.

Image feature point extraction could be affected by factors such as the information richness and feature algorithm selection, so the number of feature points varies. This poses a great burden on typical feature-matching methods. Meanwhile, the complexity of images often results in a high number of false matches, particularly in the non-rigid case. Therefore, a robust procedure for removing mismatches is desirable [5].

To address these issues, this study proposes unilateral grid-based clustering (UGC), that grids the space, clusters the match features in the grid region, and then calculates the score probability to distinguish between inliers and outliers. Because the feature matching is consistent [6], [7], the non-rigid deformation of the target can interfere with this consistency constraint, that is, the constraint effect of the whole image using consistency is common. Neighborhood characteristics in mathematical analysis show that feature point matching could preserve completeness in certain neighborhoods. Therefore, using simple grid division, the consistency of neighborhood matching can be guaranteed. At the same time, a large number of outliers present the divergence characteristics in the region with feature consistency. Using the proposed clustering statistical analysis, it is easy to distinguish between true and false matches, as shown in Fig. 1.

Research on feature matching [5], [9]–[12] implements the feature structure constraints globally or locally, thus ensuring precise correspondence. That is the classic RANSAC algorithm [13] uses the geometric information to ensure precise correspondence. However, RANSAC requires that most mismatches be excluded in advance, which is time-consuming.

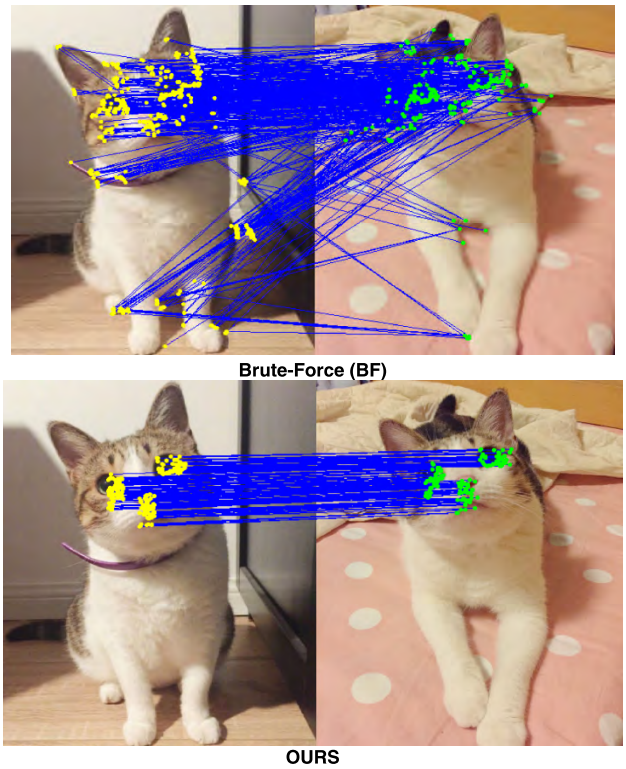


FIGURE 1. The brute-force algorithm can be well matched, but it is easily affected. Although weaker ORB descriptors are used, the proposed UGC solution can leverage feature numbers to improve the quality while maintaining real-time performance.

GMS [8] uses an efficient grid-based score estimator that can be incorporated into a real-time feature matcher, and converts the motion smoothness constraints into statistical measures for rejecting false matches. GMS distinguishes between the true or false matches by simply calculating the number of neighborhood matches. Although the theory is clear, the effect is remarkable. From a statistical standpoint, the greater the amount of data, the more realistic the statistical effect. Therefore, UGC puts forward a new direction: true matches converge to a certain extent, while false matches show divergence. Thus, the distinction between true and false matches can be made more efficiently and simpler.

Our contribution in this study includes the following:

- Put forward an efficient spatial unilateral grid. Unlike the usual matching algorithm to simultaneously grid the image pairs, it only needs to unilaterally grid one of the images in the matching pairs. This process addresses the influence of the image pairs to gridding on the statistics of the feature points, and makes our method highly efficient with real-time performance.
- Establish a method of local clustering statistical analysis. This clustering constraint utilizes the consistency between the neighborhood feature points, which is an inherent pillar of feature matching, and therefore can handle a large number of outliers.
- Prove that our UGC system has higher efficiency and more robust than traditional RANSAC [13], Fast RANSAC [14], FLANN [15], [16], LLT [5], GMS [8], and LPM [17] on standard test sets. It is important for real-time video images.

II. RELATED WORK

Feature matching has been widely applied in many fields such as target identification [18], [19], 3D reconstruction [20], motion analysis [21], and image mosaic [22]. Feature matching finds the correspondence of the effective matching points between the two images to determine the position relation of these two images. The usual strategy is a two-stage process [23]–[25]. In the first stage, a set of putative correspondence is computed using similar constraints that require the point to only be matched with similar local descriptors (such as SIFT [9], shape context [26], ORB [27], SURF [28], A-SIFT [29], Harris corners [30], and affine covariant region detectors [31]). The correspondence set of this putative not only contains most true matches, but also a large number of false matches or outliers, due to the fuzziness of the similar constraint conditions. The second stage eliminates outliers using geometric constraints that require matching to meet basic geometry requirements. The inliers and geometric parameters of the transformation are then obtained. The main challenge is how to quickly eliminate false matches and retain true ones. For example, the classic RANSAC algorithm [13], [32]–[37] can leverage geometric information to alleviate this problem. It can adopt the iterative approach to determine the optimal parameter model and eliminate the points that do not conform to the optimal model in a data set including "outlier points." However, RANSAC requires most false matches to be pre-eliminated. The Kd-tree [38], [39] establishes a match data structure by finding the most similar match pairs with image feature descriptors. In addition, some algorithms [40]–[42] first establish feature block matches and then match the corresponding feature points on the basis of block matches. The advantage of this method is that the search scope is reduced and the match precision is improved. The ICP algorithm [17], [43] uses the Euclidean distance as the match measure function through a given initial value, calculates the transform coefficient matrix, accepts the coefficient values obtained, and iteratively optimizes the objective function continuously until the target requirement is met. The disadvantage of the algorithm is that it is easy to obtain the local optimum.

Recently, grid-based motion statistics (GMS) performed local region matching with motion smoothness as a statistic. GMS enables the translation of high match numbers into high match quality, which provides a real-time, super match system. Because GMS relies on the grid to create statistics, if the grid is poorly divided or the feature points are insufficient in number, then it does not reflect the consistency of the true feature points well.

In this study, UGC shares a new design concept. A unilateral grid clustering method solves the influence of the image

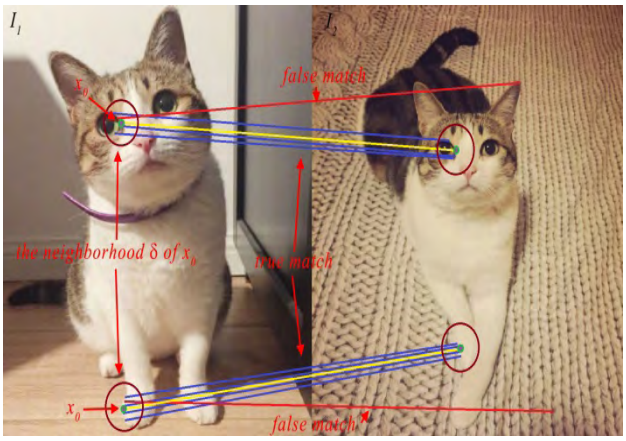


FIGURE 2. In the neighborhood of the match point, the true match points are consistent with it, while the false ones are not.

pairs to gridding on feature point statistics. At the same time, the consistency of the feature points in the local region is described. Therefore, UGC can address a large amount of high quality and high efficiency outliers.

III. PROPOSED APPROACH

In this section, we describe the proposed UGC feature-matching algorithm in detail. We start by introducing the statistics consistency phenomena of feature point data in the local region and then propose the unilateral grid division constraint. We subsequently apply the clustering approach to feature point sieving. Finally, we analyze the computational complexity, algorithm effect, and real-time performance.

A. CONSISTENCY ANALYSIS OF THE LOCAL REGION

In the same scene, the registration based on the whole image does not guarantee the consistent motion of all feature points. Some region feature points move to the left, while others move to the right. However, the consistencies of these motions are of the region. We note that other true match points should be consistent with a pair of true match points in their local neighborhood, as shown in Fig. 2.

We assume the following:

Two feature points (x_0, y_0) and (x, y) exist in neighborhood $U(x_0, \delta)$ of the feature point x_0. If feature point $x \in U(x_0, \delta)$, then the true match feature point y that corresponds to x must also be in neighborhood $U(y_0, \delta_1)$ of the true match feature point y_0 that corresponds to x_0. That is,

$$x \in U(x_0, \delta) \Rightarrow y \in U(y_0, \delta_1) \tag{1}$$

This assumption implies that the true match point pairs in the local region are consistent, while the false ones are random in the spatial position and may appear in any region in the whole image. The scene suffering from repeated textures, belongs to the special scene, and is not considered in the assumption. Therefore, the true match points in the local region should follow a similar consistency distribution rule, and the false ones should follow a random distribution without rules. This means that the local consistency distribution could serve as a useful indicator for differentiating between true and false matches.

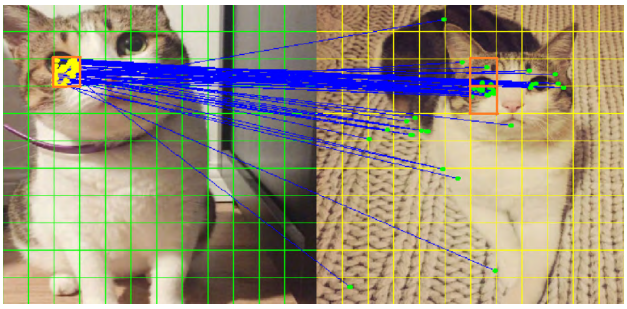


FIGURE 3. The pair of matching images is divided at the same time, the correct matching points are divided into the two adjacent grids.

B. UNILATERAL GRID DIVISION

The motion of the image feature point is smooth in a specific region scope. However, due to the influence of translation, rotation, and deformation, the consistency of feature point pairs is destroyed. This reduces the distinction between true and false match distributions. Therefore, it is very difficult to match the motion consistency of the whole image, while it is easy to obtain the smooth motion of a small enough region. If the matching pairs are simultaneously gridded, it will affect the speed of grid traversal and make the statistical results depend on the size of the mesh. As shown in Fig. 3, when a pair of matching images is divided at the same time, the correct matching points are divided into two adjacent grids (shown in the orange box), so the matching points of small proportion are removed in two adjacent grids.

Therefore, this study redesigns the idea of the unilateral grid. It only needs to divide one image into n∗n grid regions, and the other image does not need to be divided, as shown in Fig. 4. The feature points of each grid region calculate the corresponding match feature points in the match image individually. Because the true match points have consistency distribution characteristics while the false ones have random distribution characteristics, the probability distribution of the false feature match points can be expressed as

$$p_f^a = 1/(n * n) \tag{2}$$

$$p_f^{\bar{a}} = (n * n - 1)/(n * n) \tag{3}$$

where a is the grid region corresponding to the feature point and $\bar{a}$ is the region beyond the grid region corresponding to the feature point.

The distribution of feature points through the grid illustrates that the true match points in the corresponding region have a high probability, while the false ones do not, is shown in Fig. 5. Therefore, the standard deviation of the feature points can distinguish between the true and false match points well, the standard deviation of the true match points are stable, and the false ones exhibit more fluctuation. The mean and standard deviation of the feature points are, respectively:

$$\mu = \left(\sum_{i=1}^{n} d_i\right)/n \tag{4}$$

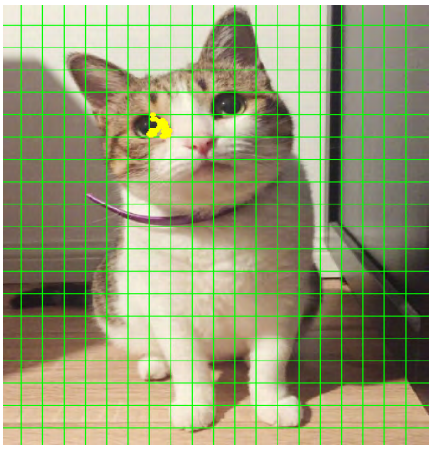


FIGURE 4. The image is divided into a 20∗20 grid region, where the yellow dots are all feature points of a grid.

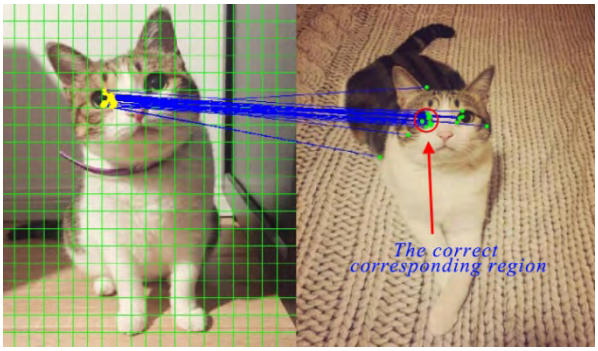


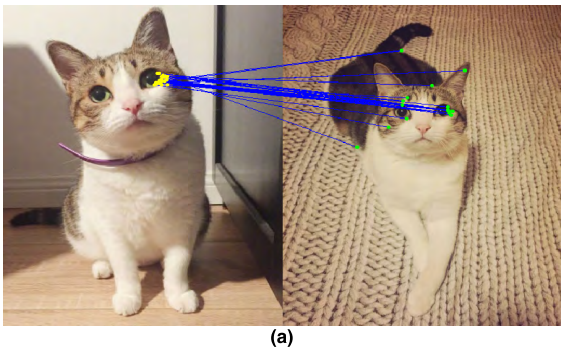
FIGURE 5. The feature points in a grid of the left figure correspond to the match points of the right figure. It is obvious that the true match points in the right figure present the aggregation characteristic (falling in the red corresponding region), while the false ones present the scattered characteristic.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \mu)}{n}} \tag{5}$$

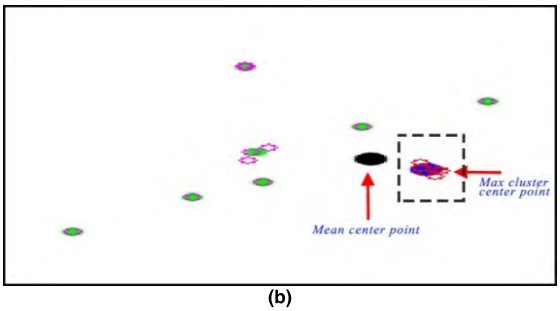
where d_i is the distance between the feature point and the center of each grid.

C. GRID CLUSTERING STATISTICS

Because of the previous clear derivation, we obtain some useful conclusions. Taking advantage of the true match points with high probability density, if the number of the true match points is greater than the false ones, then the grid region points falling into the corresponding clustering center increase continuously. The standard deviation of these grid region points is small, the data is smooth, and differentiation degree is more obvious. Even if the number of true match points is less than the false ones due to the randomness of the false match points (which presents the discrete state), the standard deviation is larger, the data are fluctuating, and it is easy to distinguish between and exclude the false match points in this region, as shown in Fig. 6. Therefore, when the true match points reach a certain proportion, the obtained match points have a very high precision and recall rate using the clustering and threshold value judgments.



(a)



(b)

FIGURE 6. a) The initial match figure of a grid region. b) The feature points of the match figure, where red represents the feature points, the black circle represents the mean center points of the feature points, green represents the clustering center points, and the blue circle represents the largest clustering center points. It is easy to determine that the blue clustering points are the true match region.

Because the mean value of the feature match points is disturbed by the false match points, we choose the clustering method to obtain the feature match center points, as shown in Fig. 6 (a mean center point and a clustering center point). Fig. 6 illustrates that the standard deviation with the higher reliability is obtained using the clustering center point.

$$\sigma_c = \sqrt{\frac{\sum_{i=1}^{n} (D_i - \mu_c)}{n}} \tag{6}$$

D_i is the distance between the clustering feature points and center points, and μ_c is the mean of the distance between the clustering feature points and center points.

We choose meanshift [44] to obtain the clustering center. Typically, the data distribution may be more than one clustering center, such as uniform distribution. There is no denseness of the points, so there many clustering centers will be present. Meanshift relies on the probability density of the cluster, and several clustering centers are obtained according to the probability density distribution, whose computation of the drift vector is as follows:

$$M(x) = \frac{\sum_{i=1}^{n} G_H\,(x - x_i)\, w(x_i)\, (x - x_i)}{\sum_{i=1}^{n} G_H\,(x - x_i)\, w(x_i)} \tag{7}$$

Here:

$$G_H\,(x - x_i) = |H|^{-1/2}\, G(H^{-1/2}(x - x_i)) \tag{8}$$

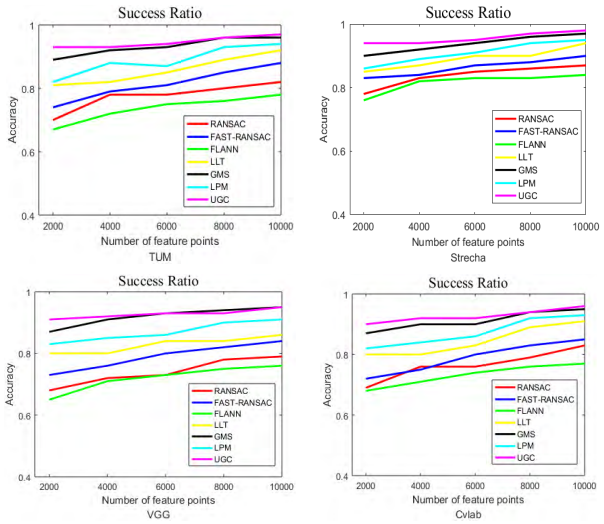


FIGURE 7. The correct rate of each algorithm with different numbers of feature points. UGC (pink) is consistently near the top. It also ensures high accuracy, especially in the case of low numbers of feature points.

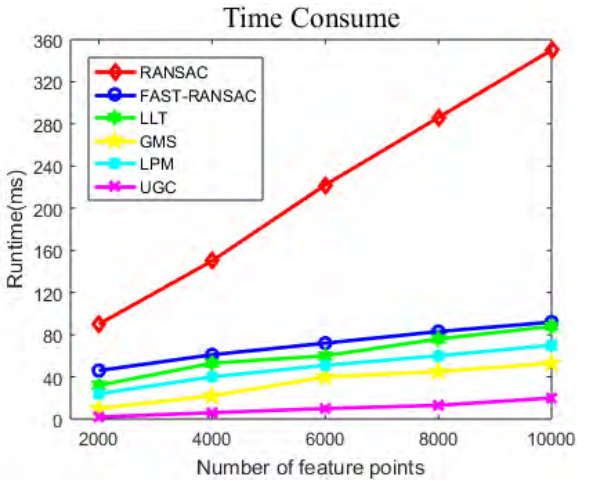


FIGURE 8. Time consumption of each algorithm with different number of feature points. UGC (pink) is consistently near the bottom. This illustrates that its runtime can reach real-time processing.

where $G(x)$ is the unit kernel function, H is the positive definite symmetric $d \times d$ matrix, and $w(x_i)$ is the weight of the sampling point x_i.

To prevent the distance of the adjacent clustering center from being too close, we adopt multiples of the grid radiuses as the combined radius R of the clustering center, judge whether the clustering centers are combined, and calculate the new combined center:

$$C_i = \frac{k_{new}C_{new} + k_i C_i}{k_{new} + k_i} \; if \; \min[d(C_{new} - C_i)] < R \qquad (9)$$

where C_i is the i_{th} clustering center point, C_{new} is the new clustering center point, k_i is the number of the i_{th} clustering points, k_{new} is the number of the new clustering points, and $d(C_{new} - C_i)$ is the distance between the new and the i_{th} clustering center point.

Since it takes much more time to calculate the standard deviation, we use effective, more concise statistical methods

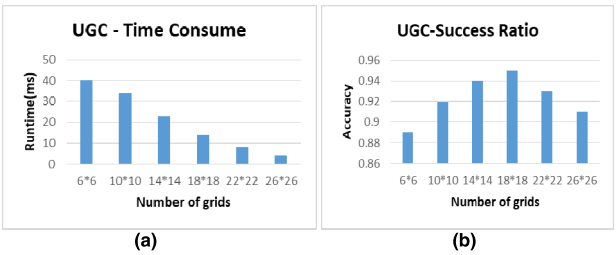


FIGURE 9. The effect of different mesh numbers on the accuracy and time consumption of the UGC. As the number of grids increases, the UGC algorithm consumes more time, and the right number rises first and then slowly declines.

to determine whether the clustering is a true or false match. The previous works typically optimized the match quality by increasing the number of feature points. However, this equally increased the algorithm operation time. In addition, meanshift clusters according to the space probability density of the feature points, without requiring a large number of feature points. Therefore, the proportion of each clustering category in the grids could be obtained by calculating the feature point numbers of each clustering category:

$$r_i = \frac{k_i}{\sum_{j=1}^{n} k_j} \qquad (10)$$

When $\max(r_i)$ is much larger, it indicates that the clustering category has higher probability density; therefore, we set a fixed threshold T. When $\max(r_i)$ is greater than the threshold T, we assume that this clustering category is the true match point set. This is of practical importance, as T is unknown and scene pair dependent. It also increases the UGC's generality by allowing it to scale with improvements in feature descriptor design.

We summarize the UGC algorithm for feature matching in Algorithm 1.

Algorithm 1 UGC Feature Matcher

Input: One pair of images
Output: Inliers set
1: Detect orb feature points
2: Use brute-force for initial matching point
3: Divide one image src into G grids
4: **for** i = 1 to G do
5: Compute the clustering centers C_j of corresponding image feature points;
6: **repeat:**
7: **if** $\min[d(C_{new} - C_j)] < R$ **then**
8: $C_j = \frac{k_{new}C_{new} + k_j C_j}{k_{new} + k_j}$
9: **end if**
10: **until:** traverse all clustering centers
11: **if** $\max(r_j) > T$ then
12: C_j is true matching points
11: **end for**

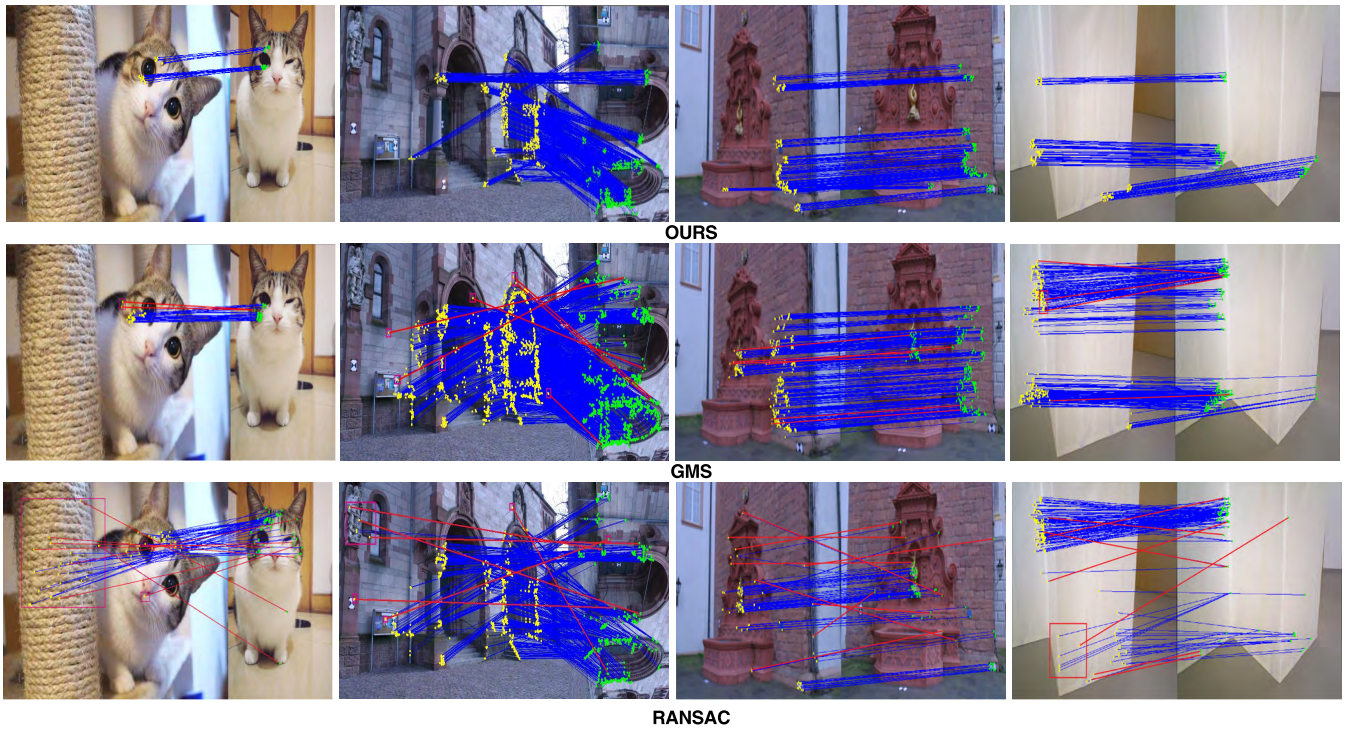


FIGURE 10. The resulting diagram of each algorithm to discard mismatches. The red boxes and red lines in the maps are partially visible mistake matching points. It is clear that UGC and GMS discard most mismatches. However, UGC matches are more accurate and support real-time, wide-baseline video matching.

IV. EXPERIMENTS

The performance of the proposed UGC algorithm was evaluated from two aspects: accuracy and time consumption. We compare UGC to powerful matchers like RANSAC, Fast-RANSAC, FLANN, LLT, GMS, and LPM. The parameters of the algorithms are consistent throughout the experiments. The experiments are conducted on a 2.4-GHZ Intel core CPU with 8-GB memory, using open source toolbox opencv3.0.

A. DATA SETS AND SETTINGS

We evaluate the performance on four data sets: TUM [45], Strecha [46], VGG [31], and Cvlab [47]. TUM contains six video sequence sets of size 640∗480, and are respectively named desk_with_person, walking_static, and sitting_static with challenging wide-baselines, low-texture, blur and rigid deformation. Strecha is a standard data set containing 500 images. These images are the 3D plaster model of size 640∗480, and they have good texture and various light intensities and angles. VGG is an affine transform data set containing 40 images, including angle change, zoom, rotation, and blur. Cvlab includes a six series sets of images, which include car, fountain, and building, with different angle motion changes and non-rigid deformation.

B. EXPERIMENTAL RESULTS

To compare the performances of each algorithm, we use the OpenCV ORB feature uniformly, with the feature numbers fixed as 1000, 3000, 5000, 8000, and 10000. The main purposes of our algorithms are to sieve the match feature points, thereby using the brute-force (BF) match algorithm to perform the initial match after collecting the feature points. The BF algorithm can use the GPU to improve the match speed. Two important parameters must be set in the UGC precision registration process: R and T. Parameter R affects the distance between the clustering categories. Parameter T judges the true clustering categories or false ones. According to the experimental test, we set R as 75% of the grid radius and T as the 50% of the total number of grid feature points.

We select the number of the feature points of difference, and compare the match sieving results of the feature points of UGC with the other algorithms in different data sets. Fig. 7 lists the accuracies of each algorithm, and the accuracy is measured by the percentage of the distance ratio. Usually, the distance between two points within 3 pixels does not affect the attitude estimation, so this requires the error to be less than 3 pixels. Fig. 7 shows that UGC is robust for the difference in the number of the feature points, and it is more accurate than the other algorithms. At the same time, the UGC algorithm also satisfies real-time performance requirements and does not include GPU acceleration. Fig. 8 shows that the average runtime of the algorithms are based on the different numbers of feature points. The runtime does not contain the time of the initial match. The data set is Cvlab. The image sizes are all normalized to 640∗480, and the grid of the algorithm is set to 18∗18, not using GPU acceleration. The UGC algorithm not only ensures high accuracy performance, but also maintains the real-time performance.

To show that the algorithm performance is affected by the different grid numbers, we compare several sets of parameters for the grid numbers, as shown in Fig. 9. The grid number is less affected by the diversity of data sets, so we only choose the Cvlab data set, and the number of feature points is 8000. When the grid numbers increase, the number of filtered feature points also decrease and the precision of the UGC algorithm shows a downward trend after reaching a certain height. This is because a too big or too small grid would influence the data consistency, and the time consumption of the UGC algorithm decreases when the grid widths increase.

We select partial match effect pictures to display, as shown in Fig. 10. The UGC and other algorithms can sieve the false match points very well. Due to some discrete correct matching points are removed by the UGC algorithm, we can see the number of the final matching points less than other algorithms, but this does not affect attitude estimation process, and it is obvious that UGC algorithm has higher true positives and relatively lower false positives than the other two algorithms.

V. CONCLUSION

UGC sieves the true or false match pairs according to the consistency principle of adjacent match points by use of statistics. This makes the algorithm theory simpler, increases the speed, and more importantly, maintains good effects. In addition, the grid selection and image quality of the UGC algorithm are related to the match performance, which will be an interesting research direction to improve the match performance.

REFERENCES

[1] J. Ma, C. Chen, C. Li, and J. Huang, "Infrared and visible image fusion via gradient transfer and total variation minimization," *Inf. Fusion*, vol. 31, pp. 100–109, Sep. 2016.

[2] A. Wong and D. A. Clausi, "ARRSI: Automatic registration of remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1483–1493, May 2007.

[3] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, no. 4, pp. 325–376, Dec. 1992.

[4] J. Ma, J. Zhao, Y. Ma, and J. Tian, "Non-rigid visible and infrared face registration via regularized Gaussian fields criterion," *Pattern Recognit.*, vol. 48, no. 3, pp. 772–784, 2015.

[5] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6469–6481, Dec. 2015.

[6] W.-Y. D. Lin, M.-M. Cheng, J. Lu, H. Yang, M. N. Do, and P. Torr, "Bilateral functions for global motion modeling," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 341–356.

[7] C. Wang, L. Wang, and L. Liu, "Density maximization for improving graph matching with its applications," *IEEE Trans. Image Process.*, vol. 24, no. 7, pp. 2110–2123, Jul. 2015.

[8] J. Bian, W.-Y. Lin, Y. Matsushita, S.-K. Yeung, T.-D. Nguyen, and M.-M. Cheng, "GMS: Grid-based motion statistics for fast, ultra-robust feature correspondence," in *Proc. IEEE CVPR*, Jul. 2017, pp. 4181–4190.

[9] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[10] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, 2011, vol. 58, no. 11, pp. 2564–2571.

[11] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 404–417, 2006.

[12] K. M. Yi, E. Trulls, V. Lepetit, and P. Fua, "LIFT: Learned invariant feature transform," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 467–483.

[13] A. M. Fischler and C. 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.

[14] Y. Zhao, R. Hong, J. Jiang, J. Wen, and H. Zhang, "Image matching by fast random sample consensus," in *Proc. Int. Conf. Internet Multimedia Comput. Service*, 2013, pp. 159–162.

[15] M. Muja and D. G. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," in *Proc. Int. Conf. Comput. Vis. Theory Appl. (VISSAPP)*, 2009, pp. 331–340.

[16] M. Muja and D. G. Lowe, "Scalable nearest neighbor algorithms for high dimensional data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 11, pp. 2227–2240, Nov. 2014.

[17] J. Ma, J. Zhao, H. Guo, J. Jiang, H. Zhou, and Y. Gao, "Locality preserving matching," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, 2017, pp. 4492–4498.

[18] J. Ma, J. Jiang, H. Zhou, J. Zhao, and X. Guo, "Guided locality preserving feature matching for remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4435–4447, Aug. 2018.

[19] J. Ma, W. Qiu, J. Zhao, Y. Ma, A. L. Yuille, and Z. Tu, "Robust L2E estimation of transformation for non-rigid registration," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1115–1129, Mar. 2015.

[20] F. Chung, "Spectral graph theory," CBMS, 1997.

[21] X. Bai, X. Yang, L. J. Latecki, W. Liu, and Z. Tu, "Learning context-sensitive shape similarity by graph transduction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 861–874, May 2010.

[22] A. A. Fathima, R. Karthik, and V. Vaidehi, "Image stitching with combined moment invariants and SIFT features," *Procedia Comput. Sci.*, vol. 19, pp. 420–427, Dec. 2013.

[23] J. Ma, J. Zhao, J. Tian, A. L. Yuille, and Z. Tu, "Robust point matching via vector field consensus," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1706–1721, Apr. 2014.

[24] S. Pang, J. Xue, Q. Tian, and N. Zheng, "Exploiting local linear geometric structure for identifying correct matches," *Comput. Vis. Image Understand.*, vol. 128, pp. 51–64, Nov. 2014.

[25] J. Ma, J. Jiang, C. Liu, and Y. Li, "Feature guided Gaussian mixture model with semi-supervised EM and local geometric constraint for retinal image registration," *Inf. Sci.*, vol. 417, pp. 128–142, Nov. 2017.

[26] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 509–522, Apr. 2002.

[27] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571.

[28] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vis.*, 2006, pp. 404–417.

[29] J.-M. Morel and G. Yu, "ASIFT: A new framework for fully affine invariant image comparison," *SIAM J. Imag. Sci.*, vol. 2, no. 2, pp. 438–469, 2009.

[30] K. G. Derpanis, "The Harris corner detector," in *Proc. Symp. Svenska Sllskapet Bildanalys*, 2004.

[31] K. Mikolajczyk *et al.*, "A comparison of affine region detectors," *Int. J. Comput. Vis.*, vol. 65, no. 1, pp. 43–72, 2005.

[32] P. H. S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Comput. Vis. Image Understand.*, vol. 78, no. 1, pp. 138–156, 2000.

[33] T.-J. Chin, J. Yu, and D. Suter, "Accelerated hypothesis generation for multistructure data via preference analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 625–638, Apr. 2012.

[34] R. Raguram, J.-M. Frahm, and M. Pollefeys, "A comparative analysis of RANSAC techniques leading to adaptive real-time random sample consensus," in *Proc. Eur. Conf. Comput. Vis.*, 2008, pp. 500–513.

[35] O. Chum, J. Matas, and J. Kittler, "Locally optimized RANSAC," in *Proc. Joint Pattern Recognit. Symp.*, 2003, pp. 236–243.

[36] T. Sattler, B. Leibe, and L. Kobbelt, "SCRAMSAC: Improving RANSAC's efficiency with a spatial consistency filter," in *Proc. Int. Conf. Comput. Vis.*, 2009, pp. 2090–2097.

[37] M. Kushnir and I. Shimshoni, "Epipolar geometry estimation for urban scenes with repetitive structures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 12, pp. 2381–2395, Dec. 2014.

[38] Y.-S. Kang, J.-H. Nah, W.-C. Park, and S.-B. Yang, "gkDtree: A group-based parallel update kd-tree for interactive ray tracing," *J. Syst. Archit.*, vol. 59, no. 3, pp. 166–175, 2013.

[39] L. Q. Liao, J. S. Bai, and D. A. Luo, "Integrated point cloud storage structure based on octree and KDTree," *Comput. Syst. Appl.*, vol. 21, no. 3, pp. 87–90, 2012.
[40] J. Xie and P.-A. Heng, "Shape modeling using automatic landmarking," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 8, no. 2, 2005, pp. 709–716.
[41] A. Myronenko, X. Song, and M. Carreira-Perpinán, "Non-rigid point set registration: Coherent point drift," in *Proc. Neural Inf. Process. Syst.*, vol. 19, 2007, pp. 1009–1016.
[42] A. Toshev, J. Shi, and K. Daniilidis, "Image matching via saliency region correspondences," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
[43] Y. Liu, "Improving ICP with easy implementation for free-form surface matching," *Pattern Recognit.*, vol. 37, no. 2, pp. 211–226, 2004.
[44] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.
[45] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems," in *Proc. Int. Conf. Intell. Robot Syst. (IROS)*, Oct. 2012, pp. 573–580.
[46] C. Strecha, W. von Hansen, L. Van Gool, P. Fua, and U. Thoennessen, "On benchmarking camera calibration and multi-view stereo for high resolution imagery," in *Proc. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2008, pp. 1–8.
[47] K. M. Yi, Y. Verdie, P. Fua, and V. Lepetit, "Learning to assign orientations to feature points," in *Proc. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 107–116.

ZHAOHUI ZHENG received the B.S degree in mathematics and applied mathematics from the Huanggang Normal College and Technology in 2009 and the M.S degree in applied mathematics from the Wuhan University of Technology, Wuhan, China, in 2011. He is currently pursuing the Ph.D. degree with the School of Electronic Information, Wuhan University, Wuhan. His research interests include machine learning, pattern recognition, robotic dynamics, and mathematical modeling.

YONG MA received the B.S. and M.S. degrees from the Department of Automatic Control, Beijing Institute of Technology, Beijing, China, in 1997, and the Ph.D. degree from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2003. From 2004 to 2006, he was a Lecturer at the University of the West of England, Bristol, U.K. From 2006 to 2014, he was with the Wuhan National Laboratory for Optoelectronics, HUST, where he was a Professor of electronics. He is currently a Professor with the Electronic Information School, Wuhan University. His general field of research is in signal and systems. His current research projects include remote sensing of the Lidar and infrared, infrared image processing, pattern recognition, and interface circuits to sensors and actuators.

HONG ZHENG received the Ph.D. degree from Wuhan University. He went to Deakin University for post-doctoral research in 2001, Australia. He is currently a Professor with the Electronic Information School, Wuhan University. His field of research is machine vision, image processing, biometrics, intelligent systems, and robotic dynamics.

JIANPING JU received the B.S. degree in optical information science and technology from the Huazhong University of Science and Technology and the M.S. degree in artificial intelligence and pattern recognition from the Wuhan Institute of Technology, Wuhan, China, in 2003 and 2010, respectively. He is currently pursuing the Ph.D. degree with the School of Electronic Information, Wuhan University, Wuhan. His research interests include machine learning and pattern recognition.

MINGYU LIN received the master's degree in signal and information processing from Nanchang Hangkong University. She is currently a Lecturer with the School of Electrical and Information Engineering, Hubei Business College. Her filed of research is image processing and pattern recognition.