IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 30, 2020, accepted September 8, 2020, date of publication October 12, 2020, date of current version October 28, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3030183

Feature Detection and Matching With Linear Adjustment and Adaptive Thresholding

ZHIMING CAI[1,2], YIWEN OU[1], YUFENG LING[1], JIAN DONG[1], JIAN LU[1], AND HOWARD LEE[3]

[1]School of Information Science and Engineering, Fujian University of Technology, Fuzhou 350118, China
[2]National Demonstration Center for Experimental Electronic Information and Electrical Technology Education, Fujian University of Technology, Fuzhou 350118, China
[3]Leedwell Pty. Ltd., Melbourne, VIC 3141, Australia

Corresponding author: Zhiming Cai (caizm@163.com)

This work was supported in part by the Postgraduate Teaching Reform Project Teaching Case Base of Machine Recognition and Image Processing, and in part by the New Engineering Research Project Construction of Electronic and Electrical Practice Education System and Practice Platform Oriented to New Engineering of Fujian University of Technology.

ABSTRACT Image feature detection and matching technologies are crucial aspects in machine vision. However it is still facing the dilemma between fast operation for real-time application and robust matching. To address this issue, we propose a robust and relatively fast method for image feature extraction and matching with linear adjustment and adaptive thresholding (LAAT) in this paper. The major challenge of this method is reducing the sensitivity to the brightness. To solve this problem, we adopt brightness and contrast adjustment for image pairs processed by Gaussian filtering. An adaptive thresholding FAST approach is applied for feature selection to improve the performance. The proposed method is compared with the traditional and state-of-the-art extraction methods on public dataset. Particularly, this paper focuses on the illumination change, image blur, and image rotation aspects. Experiments show that our proposed algorithm is superior to other algorithms in the comprehensive evaluation of various parameters, especially for illumination and blur transformations.

INDEX TERMS Adaptive thresholding, contrast adjustment, feature detection and matching, robustness.

I. INTRODUCTION

Machine vision is a branch of artificial intelligence which is developing rapidly. In short, machine vision uses cameras instead of human eyes to measure and judge surrounding environment. Feature detection and matching play important roles in machine vision. They are widely used in many fields, such as 3D reconstruction{Mouragnon, 2006 #158}., posture estimation [2], Smart device application [3], AR [4], SLAM (simultaneous localization and mapping) [5], [6], robot navigation [7], object recognition [8], etc.

Image feature detection and matching algorithms consist of three main steps: (1)detect feature and compute descriptors; (2)match descriptor; (3)remove false matches. First, local feature detector is used to find feature points, and then feature descriptors are used to describe them in a compact manner. Second, descriptors are matched with Euclidean distance or Hamming distance. Third, in order to improve the accuracy and the robustness in the process of descriptor matching, false matches need to be removed as error matching inevitably exists.

There are many different approaches for feature detection, descriptor computing and image matching, however a large performance gap exists between much faster but often unstable real-time solutions and slow but robust matching methods. For example, the features detected by SIFT [9] algorithm not only have strong robustness for the illumination, scale and rotation transformation, but also can maintain certain stability for noise and perspective transformation too. However, SIFT suffers a huge computational complexity yet.

The SURF [10] algorithm is improved on the basis of SIFT algorithm. It uses the determinant of Hessian matrix to determine the position of interest points, and then determines the descriptors according to the Haar wavelet response of the neighborhood points of interest. Thus, in addition to the high repeatability of detection and the good distinguish ability of descriptors, SURF algorithm also has strong robustness and higher operation speed compared with SIFT. As presented in the previous study [10], SURF algorithm is more than several times faster than SIFT algorithm. Although the

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

comprehensive performance of SURF algorithm is better than SIFT algorithm, it cannot achieve real-time performance.

In 2011, Rublee *et al.* proposed ORB [11] (oriented FAST and rotated binary), which is a FAST algorithm for feature extraction and description. ORB uses FAST [12] algorithm to detect feature points. The most outstanding advantage of FAST algorithm is its computational efficiency, which is faster than other classical feature point extraction methods (such as SIFT, Susan, Harris). Although it takes less time, ORB, which depending much on a fixed threshold value t, normally t=*20*, has less variance to geometric or photoelectric change and it is easily affected by image noise.

To address this issue, there are many studies that had improved in some deficiencies. Mur-Artal *et al.* presented an improved ORB algorithm and applied it to SLAM (ORB-SLAM) [5]. The result had improved the uniformity of feature distribution. They divided each layer of image into grids, and extracted at least 5 feature points in each grid. When the number of feature points is insufficient, the threshold will be adjusted accordingly. Cao *et al.* [13] proposed a fast and robust local feature extraction method, called OOD, which has high performance in both computational cost and speed. Kuang *et al.* [14] optimized the matching strategy based on the premise of keeping the matching accuracy unchanged. The improved KD tree was used to optimize the matching strategy of ORB algorithm and the matching speed had been improved. Fausto *et al.* [15] proposed a new feature descriptor called spider local image features (SLIF),which built the feature vectors by selectively extracting image information from a group of previously detected feature point. This method produces simple low-dimensional feature descriptors and is robust to several image transformation and distortions, such as viewpoint, scale and rotation transformations. Some studies were aimed to improve in two aspects: generating descriptors and eliminating error matches. Huang *et al.* [16] presented a fast and robust feature matching method with binary affine invariant descriptor (improves the affine invariance of binary descriptor with fast processing) and local geometric consistency check (false matches are efficiently filtered out to achieve high recall). Li *et al.* [17] had improved a method both in feature detection and description. In order to improve the scale invariance, the proposed method built a pyramid-like scale space, to detect OFAST feature points on each layer. They had utilized the 128-bit improved FREAK description operator to replace the RBRIEF description operator that had the last 128 bits of the small variance, which improved the matching accuracy and robustness. Wang *et al.* [18] proposed a novel feature extraction algorithm named GA-SURF, which is based on SURF and the theory of Geometric Algebra (GA) to process multispectral images. The GA-SURF algorithm has both fast speed in computation and strong robustness in recognition. However, compared with SURF algorithm, GA-SURF has a worse performance in the rotation and JPEG compression transformation.Other literatures adopted the merits of combining multiple algorithms to improve the performance. L-SURB [19] used SURF to detect feature points and ORB to describe feature points, after the image had been enhanced by Laplacian operator. The algorithm effectively solves the problem in ORB algorithm that is sensitive to image brightness and poor in scale invariance, which greatly improves the matching accuracy. Li *et al.* [20] developed a fast matching algorithm based on the combination of FAST feature points and SURF descriptor. Their work had improved the matching effect and real-time property for objects with rich feature points. Cheng *et al.* [21] presented the algorithm that combined the SURF with Harris in the image matching process. The proposed algorithm improved the quality of UAV (Unmanned Aerial Vehicle) image matching with strong robustness and high efficiency. And V. A *et al.* [22] proffered two novel detector-descriptor variants: one was SURF detector with SIFT descriptor and another was SIFT detector with SURF descriptor, which were used to augment the performance of contemporary FR systems. The experimental results demonstrated that the SURF detector with SIFT descriptor method outperformed the others methods, as it could detect more number of feature points and keep robust even under unconstrained scenarios. In this paper, we focus on improving the robustness of illumination invariance and blur invariance of the feature detection and matching. And like the approach by Li *et al.* [23], we combine ORB and SURF algorithm (ORB+SURF) as one of the comparison algorithms, where ORB is used to detect feature points and SURF is used to generate descriptors for feature points. Firstly, we improve the robustness of illumination invariance and blur invariance by using Gaussian filter and linear brightness adjustment before feature detection. Secondly, we update the FAST algorithm to detect feature points by substituting an adaptive threshold with fixed threshold. Thirdly, we use the SURF method to describe the feature points. Fourthly, we utilize the simplest feature matching method, Brute-Force Matcher, to calculate the Euclidean distance between two feature point descriptors in the matching image pairs respectively. Then we sort the distances and take the nearest ones as the matching points. In this way, many mismatches still occur though the matching complexity is reduced. Therefore, we have applied some mechanisms to remove the mismatches in the last step. We use coarse matching to filter out mismatches at first, and follow by applying RANSAC [24] principle to filter out the rest of mismatches.

The structure of this paper is summarized as follows: The introduction of the relevant work and our approach is described in Section I. In Section II, we illustrate our method in detail. First, the image pairs are pre-processed by Gaussian filtering and brightness adjustment. Second, the adaptive threshold FAST algorithm is proposed to extract features. Then, we introduce the generation of feature point descriptors and match descriptors, respectively. Finally, two steps are applied to remove false matches. In Section III, we present the experimental results to evaluate our proposed method and compare with other algorithms on popular datasets.

II. OVERVIEW OF THE PROPOSED METHOD

A. THE WORKFLOW OF THE PROPOSED METHOD

In this paper, a robust and relatively fast method for image feature extraction and matching with linear adjustment and adaptive thresholding (LAAT) is proposed. The workflow of the proposed algorithm is presented in Fig. 1.

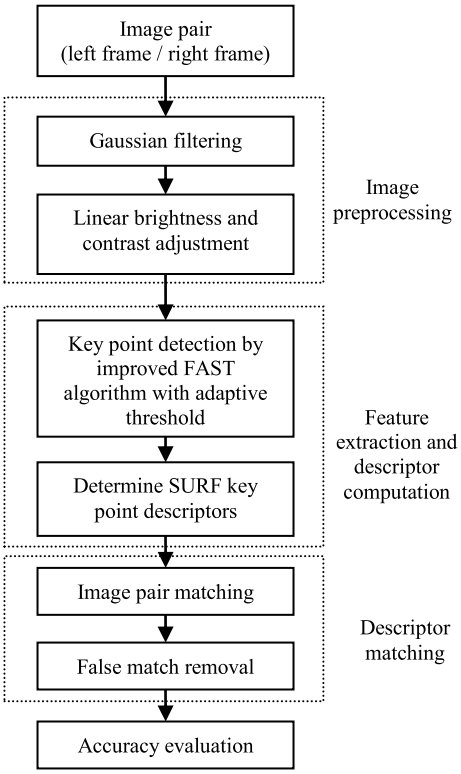


FIGURE 1. The workflow of our proposed algorithm.

B. GAUSSIAN FILTERING

Gaussian filter is a kind of linear smoothing filter, which is widely used in image processing.

The operation of Gaussian filtering is using a template to scan every pixel in an image, and using the weighted average gray value of the pixels in the neighborhood determined by the template to replace the value of the central pixel of the template.

For image processing, two-dimensional zero mean Gaussian function is as follows:

$$G_0(x, y) = A \times \exp(-(\frac{(x - x_0)^2}{2\sigma_x^2} + \frac{(y - y_0)^2}{2\sigma_y^2})) \quad (1)$$

where, A is the amplitude, (x, y) and (x_0, y_0) represent the pixel and central pixel in the template, respectively; σ_x^2 and σ_y^2 are the variances.

C. IMAGE CONTRAST, BRIGHTNESS VALUE ADJUSTMENT

In the process of image matching, due to the influence of illumination, the matching accuracy tends to decrease, so we consider using one brighter image to compensate the brightness of another. To improve the robustness of illumination invariance, this paper uses point operation to adjust the brightness and contrast of the image.

The linear brightness and contrast adjustment functions are as follows:

$$g(x, y) = a \times f(x, y) + b \quad (2)$$

$$a = \frac{mean(img1)}{mean(imgn)} \quad (3)$$

$$b = mean(img1) - mean(imgn) \quad (4)$$

where, a is gain, which needs to be greater than 0; b is bias. They are used to control the contrast and the brightness of the images, respectively. The $f(x, y)$ and $g(x, y)$ are the pixels of the input and output color images; *mean(img1)* and *mean(imgn)* are the means of the pixel values of the selected image and each remaining image in public datasets, respectively. Algorithm 1 shows the contrast and brightness adjustment for the image.

Algorithm 1 Image Contrast, Brightness Adjustment

Input: selected image *img1*, and another remaining image, *imgn*, in public datasets

Output: a contrast, brightness improved image.

Step 1 Convert *img1* and *imgn* to grayscale image.

Step 2 Calculate the mean value of these two grayscale images respectively.

Step 3 Compare these two mean values.

if $mean(img1) > mean(imgn)$

$$a = \frac{mean(img1)}{mean(imgn)}$$
$$b = mean(img1) - mean(imgn);$$

else

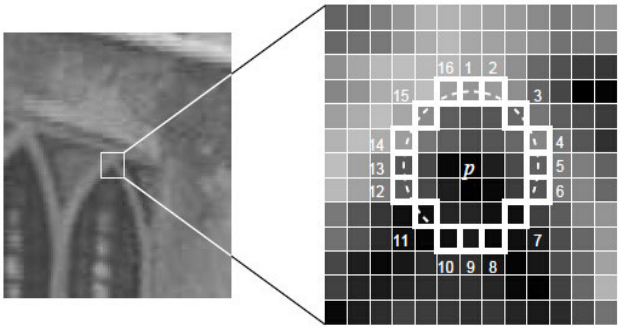
$$a = \frac{mean(imgn)}{mean(img1)}$$
$$b = mean(imgn) - mean(img1)$$

Step 4 Visit each pixel in the darker color image, and substitute a and b values calculated above into $g(x, y) = a \times f(x, y) + b$, and then

Step 5 Get the contrast, brightness improved image.

D. AN IMPROVED FAST ALGORITHM WITH ADAPTIVE THRESHOLDING

The idea of FAST detector is to compare candidate point with 16 pixels around the point, as shown in Fig. 2. In Fig. 2(a) a representative candidate point, which is located in the intersection of two lines, is selected. However, candidate points are not necessarily real corners. As illustrated in Fig. 2(b), the candidate point, whose pixel value is I_p and the surrounding sixteen pixels with a radius of 3 centered on the candidate point. The 16 pixels are denoted as p1, p2, . . . , p16.



(a) A candidate corner (b) The 16 pixels surrounding the candidate corner p.

FIGURE 2. FAST detector.

The FAST detector divides the points into three groups:

$$S_{pk} = \begin{cases} darker, & I_{pk} \le I_p - t \\ similar, & I_p - t < I_{pk} < I_p + t \\ brighter, & I_{pk} \ge I_p + t \end{cases} \quad (5)$$

where, t is a fixed threshold, I_{pk} $(k = 1, 2, 3 \ldots, 16)$ is the intensity of pixel k around p. When $I_{pk} \le I_p - t$, S_{pk} belongs to the *darker* group; when $I_p - t < I_{pk} < I_p + t$, S_{pk} belongs to the *similar* group; when $I_{pk} \ge I_p + t$, S_{pk} belongs to the *brighter* group. In this way, a circular region can be divided into three parts: *darker*, *similar* and *brighter*. As long as the number of *darker* or *brighter* in the circular area is greater than constant, r (generally r=9), then the point is considered as a corner.

In order to accelerate the speed, FAST algorithm proposes that it is not necessary to compare these pixels one by one. To simplify the process, first, compare the pixel values of points 1, 5, 9 and 13 (i.e. 4 points in the horizontal and vertical directions) with the central pixel value. If three or more of the four pixel values are greater than $I_p + t$ or less than $I_p - t$, then the central point is considered to be a candidate corner (also called feature point), otherwise it cannot be a corner.

FAST corner detection algorithm is a feature extraction operation with high computational efficiency and high repeatability. It has been widely used in stereo image matching, image registration [25], target detection [26],target recognition [27], target tracking [28] and other fields, and has become the most popular corner detection method in the field of computer vision. However, the influence of noise and the threshold have great impact on this method. Due to the FAST algorithm sets a fixed threshold value, the feature points of an image will suffer from extraction errors while the brightness changes in an outdoor environment. It is obvious that the deficiency of FAST algorithm also depends on the selection of threshold value, t, therefore it is important to determine an adaptive threshold value to replace the fixed value, t. In this paper, we propose an adaptive calculation method for threshold. We divide an image into $S \times T$ blocks. The adaptive threshold, denoted as t', is expressed as:

$$t' = \frac{\sum_{i=1}^{S \times T} M_i}{\sum_{j=1}^{S \times T} D_j} * (\sum_{i=1}^{S \times T} M_i / M_a) \quad (6)$$

where

$$M_a = (\sum_{i=1}^{S \times T} M_i - M_{\max} - M_{\min})/(S \times T - 2) \quad (7)$$

M_i denotes the mean intensity of block i, D_j is the standard deviation intensity of block j. Let B_{max}, B_{min} be the two blocks whose mean intensities are the maximum and the minimum among all blocks. M_{max} and M_{min} are the mean intensities of B_{max} and B_{min} respectively. M_a is the average of mean intensities of the blocks except B_{max} and B_{min}. The details of the adaptive threshold based FAST algorithm are provided in algorithm 2.

Algorithm 2 Improved FAST Algorithm Based on Adaptive Threshold

Input: image
Output: threshold, t'

Step 1 Divide an image into $S \times T$ blocks
Step 2 Calculate the mean and standard deviation of every block, M_i, D_j $(i, j = 1, 2, 3, \ldots, S \times T)$.
Step 3 Find the two blocks, B_{max}, B_{min}, and calculate the mean intensities, M_{max} and M_{min}.
Step 4 M_a is evaluated by using (7).
Step 5 M_a is substitute into (6), to obtain the adaptive threshold t'.
Step 6 Similar to FAST, we use $I_p + t'$ and $I_p - t'$, to evaluate the corners.

In rare cases, if insufficient feature points are detected by the improved algorithm, in order to increase the number of feature points detected, we will adjust the threshold $t' = 20$.

As the feature points detected by FAST algorithm do not have scale and rotation invariance, ORB algorithm builds image pyramid to get scale invariance, and uses the concept of intensity centroid method [29] to achieve rotation invariance.

The intensity centroid method assumes that there is an offset between the gray level of a feature point and the centroid in a small image patch. The main direction of the feature point is evaluated by the vector from the feature point to the centroid. The centroid is calculated by moment, which is defined as follows:

$$m_{pq} = \sum_{x,y} x^p y^q I(x, y) \quad p, q = \{0, 1\} \quad (8)$$

where $I(x,y)$ denotes the gray value for the pixel (x,y). The centroid C of a field is:

$$C = (\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}) \quad (9)$$

The orientation angle of the feature point is:

$$\theta = \arctan(\frac{m_{01}}{m_{10}}) \tag{10}$$

In order to further improve the robustness of rotation invariance of this method, normally, circular area is selected as the patch with the feature point as the center.

E. GENERATION OF FEATURE POINT DESCRIPTORS

We use SURF algorithm to generate descriptors for feature points.

The first step involves constructing a square region with the size of 20×20 square pixel, centered at the feature point, and oriented along the main orientation chosen in the previous part. The examples of such square regions are depicted in Fig. 3.

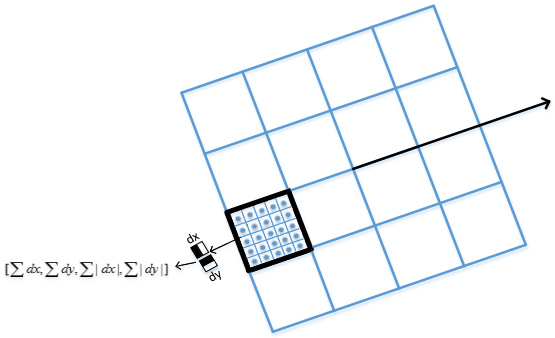


FIGURE 3. Descriptor representation.

The regions are divided into 4×4 sub-regions, and for each sub-region, a few simple feature points at 5×5 regularly spaced sample points are computed. The dx and dy represent the Haar wavelet responses in the horizontal and vertical direction (filter size of 2s), respectively. At the same time, the responses of dx and dy are weighted with Gaussian ($\sigma = 3.3$s) centered on the feature point.

Then, the sum of the wavelet responses dx and dyover each sub-region, and the sum of the absolute values of the responses $|dx|$ and $|dy|$ form a feature vector which is as follows:

$$v_{subregion} = [\sum dx, \sum dy, \sum |dx|, \sum |dy|] \tag{11}$$

Therefore, each sub-region carries a four-dimensional vector that describes the feature points. Then a 4×4 sub-region is described by a vector with the length of 64.

F. DESCRIPTOR MATCHING

We use Brute-Force Matcher to match descriptors which Euclidean distance is chosen as measure metric.

The calculation of similarity measure of two feature point descriptors is as follows:

$$Dist_{ij} = [\sum_{k=1}^{K} (X_{ik} - X_{jk})^2]^{1/2} \tag{12}$$

where, K is the dimension of the feature vector, X_{ik} represents the k^{th} element of the i^{th} feature descriptor in the image to be matched, X_{jk} refers to the k^{th} element of the j^{th} feature descriptor in the reference image.

G. DELETION OF FALSE MATHCES

In this paper two steps are applied to remove false matches:

Step1(coarse matching): Determine the minimum distance between all matching pairs. When the match distance is more than twice of the minimum distance, the match will be deleted.

Step2(fine matching): We further use RANSAC to filter mismatches. RANSAC can find the optimal parameter model in a group of data sets containing ''outliers'' by means of continuous iteration. The ''outliers'' which do not conform to the optimal model will be deleted too.

III. EXPERIMENTAL RESULTS

A. DATASETS AND EVALUATION PROTOCOLS

Experimental setup: OPENCV 3.4 has been used to perform the experiments presented in this paper. Specifications of the computer system used are: Intel®Core^{TM}i5-4210U CPU @1.70GHz 2.40GHz and 4.00GB RAM.

Parameter: we set the image division block parameters $S \times T$ to 3×4.

Datasets: We use illumination change (the Leuven sets), image blur (the bikes sets), image rotation (the boat sets) in VGG Dataset (http://www.robots.ox.ac.uk/∼vgg). Each of the sets contains 6 images with an increasing amount of photometric or geometric distortion, as shown in the Fig. 4.

(a)Leuven Dataset (b)Bike Dataset (c)Boat Dataset

FIGURE 4. The first image of each dataset. (a)Leuven dataset: The brightness of the six pictures decreases gradually. (b)Bike dataset: The ambiguity of the six images increases gradually. (c)Six images are gradually rotated and scaled.

Evaluation protocols: Mikolajczyk and Schmid [30] had proposed the evaluation criteria of feature point detection and matching. The concept of repeatability is often used to evaluate the performance of the feature point detection method. The formula is as follows:

$$Repeatability = \frac{\sum Dist(m'_{points}, n_{points})}{Min(m'_{points}, n'_{points})} \tag{13}$$

Let H be the homography matrix of image transformation and m_{points}, n_{points} be the feature point sets detected in the two images for matching. The feature point set m_{points} of first image is multiplied by H and the resultant feature points beyond the coordinates of second image are deleted. The remaining feature points are denoted as m'_{points}.

Similarly, the remaining feature points left in second image after inverse homography transformation are denoted as n'_{points}.

The definition of repeatability is extended in [31], aiming at affine invariance. The repeatability is redefined as:

$$Repeatability = \frac{\sum Correspondence}{Min(m'_{points}, n'_{points})} \tag{14}$$

where *Correspondence* denotes the repetition points, which needs to satisfy two conditions: 1)The threshold value of distance between m'_{points} and n'_{points} is less than the set value, $\xi = 1.5$ pixel. The overlap error of feature point area mapping to another image is less than 20% (the default parameter).

Another evaluation of the feature point matching is accuracy, which represents how much of the matching results are accurate. The formula is as follows:

$$Accuracy = \frac{a}{b} \times 100\% \tag{15}$$

where a is the number of interior points (defined in Algorithm 3), b is the number of features after fine matching.

Algorithm 3 Calculation of Accuracy

Input: image pairs
Output: accuracy

Step 1 Using RANSAC to calculate the homograph matrix with image pairs.
Step 2 Using homograph matrix to map point (x, y) in the first image to an estimated point (x', y') in the second image. Define

$$dx = x' - x, dy = y' - y.$$

Step 3 If $dx^2 + dy^2 < 9$, then (x', y') denotes the interior point.
Step 4 Using (15) to calculate the accuracy.

B. COMPARISON BETWEEN IMPROVED ALGORITHM AND OTHER ALGORITHMS

In this subsection, the extraction time (the process of feature point detection and descriptor computation), the number of feature points, accuracy and repeatability are chosen as the evaluation criterions. We use the above datasets to compare the algorithms (traditional ORB algorithm, ORB-SLAM, ORB+SURF and the improved algorithm LAAT), and match the first picture in each dataset with the other five pictures in turn. We take the average of five extraction time as the final result.

1) ILLUMINATION CHANGE

We match the first image from the light change dataset with the remaining five images, where the brightness of the six images decreases gradually, as indicated in the horizontal axis of Fig. 5.

The repetition rate point-line graph of the image brightness change is shown in Fig. 5(a). It is obvious that LAAT surpasses other algorithms in repetition rate. Although the repetition rate of LAAT descends as the image brightness decreases, the repetition rate still remains above 64%. With the decrease of brightness, the repetition rate of the traditional ORB and ORB-SLAM decreases sharply to 23%, and 48%, respectively. And the lowest repetition rate of ORB+SURF algorithm reaches 30%.

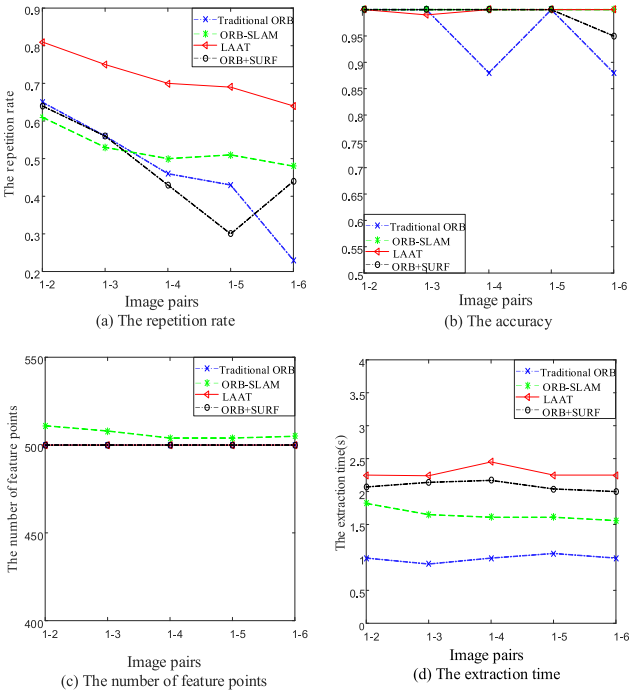


FIGURE 5. Extraction performance under varying brightness levels in the Leuven Dataset.

Fig. 5(b) shows the accuracy point-line graph of the image brightness change. The results from ORB-SLAM have impressively shown perfect accuracy, whereas, the results from LAAT also close to perfect. Although the accuracy of ORB+SURF is close to LAAT's, it drops down when the first image is matched with the darkest image. The accuracy of traditional ORB is relatively unstable. According to the Fig. 5(c), despite the changes in brightness, the number of feature points extracted by LAAT remains constant at 500. And other algorithms also keep around 500.We can see that the feature points extracted by LAAT are relatively stable.

As shown in Fig. 5(d), the extraction time of ORB is superior to the others. Although compared with other algorithms, LAAT takes longer time in extraction, however it still maintain under 2.45 seconds which nearly fulfills the basic requirement for real-time applications.

2) BLUR CHANGE

We match the first image from the blur change dataset with the remaining five images, showing increased in blur effect, as indicated in the horizontal axis of Fig. 6.

As presented in the Fig. 6(a), with the increase of blur, the repetition rate of ORB-SLAM algorithm keeps below 46%.

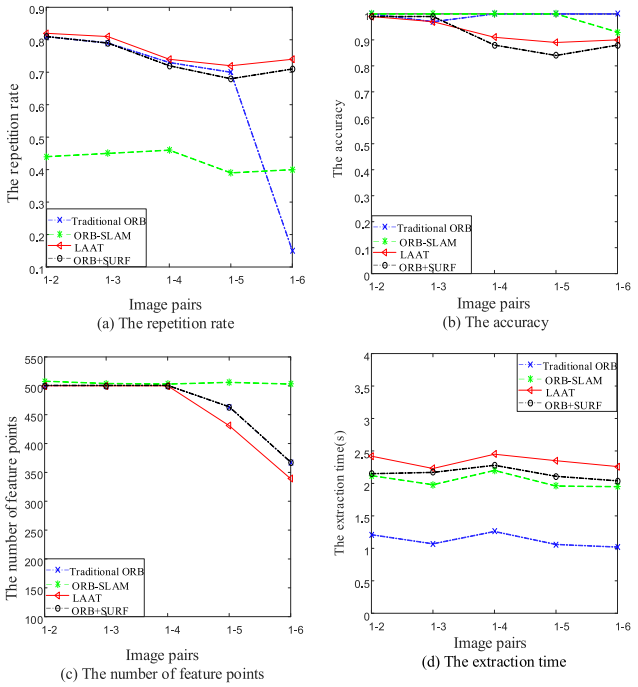


FIGURE 6. Extraction performance under varying blur levels in the Bike Dataset.

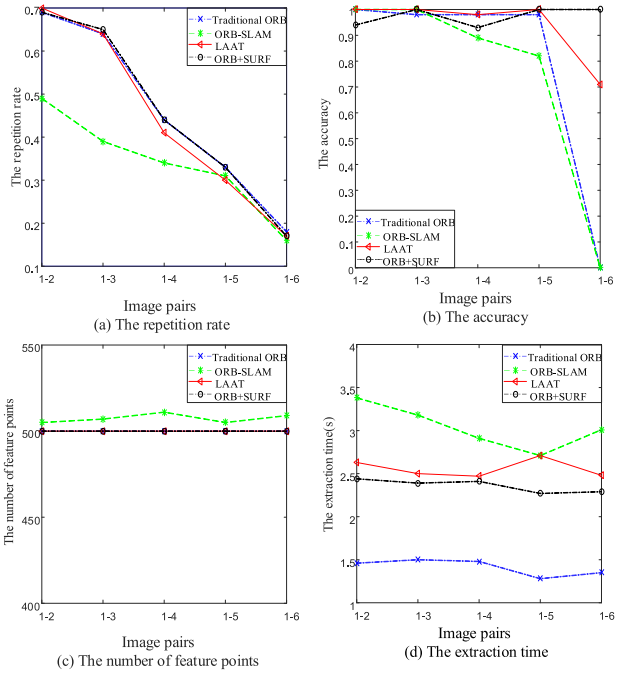


FIGURE 7. Extraction performance under varying rotation and scale levels in the Boat Dataset.

Compare to those of the traditional ORB and ORB+SURF drop from 81% to 15% and 71%, respectively. For LAAT, the repetition rate keeps above 72%, which is always higher than other three algorithms.

The accuracy vs image blur change line graph is shown in Fig. 6(b). With the change in blur, the accuracy of LAAT is slightly higher than that of ORB+SURF. Although the accuracy of LAAT is slightly lower than those of the traditional ORB and ORB-SLAM, the accuracy of LAAT mostly keeps above 90%.

From Fig. 6(c), we can see that the number of feature points of ORB-SLAM is stable over 500 while those of traditional ORB, LAAT, ORB+SURF drop down. Though the number of feature points of LAAT is the lowest one, the matching pairs obtained by LAAT are more than those of others as shown in Fig. 14(b). As shown in Fig. 6(d), the experimental results are similar to Fig. 5(d).

3) ZOOM AND ROTATION CHANGE

We repeat the matching operations in the zoom and rotation change dataset. We match the first image in the zoom and rotation change dataset with the remaining five images one by one, where the images are gradually rotated and scaled.The comparison results are depicted in the Fig. 7. As shown in Fig. 7(a), the repetition rates of all algorithms drop rapidly as the images continue to rotate and zoom. Fig. 7(a) shows the trend of the repetition rate of LAAT is still similar to those of traditional ORB and ORB+SURF. The repetition rate of LAAT is higher than that of ORB-SLAM.

According to Fig. 7(b), the accuracies of all algorithms decrease as well while the images continue to rotate and zoom. Nevertheless, the accuracy of the improved algorithm LAAT normally is higher than ORB and ORB-SLAM, which fail to find the inner points and the accuracies drop to zeros at last matches.

As shown in Fig. 7(c), the experimental results are similar to Fig. 5(c). The numbers of feature points of all algorithms are close to each other and range between 500 and 511. Fig. 7(d) shows the extraction time for image rotation and zoom changes. The extraction time of LAAT is lower than that of ORB-SLAM and close to that of ORB+SURF, which can also basically meet the requirement for real-time applications.

In summary, the repetition rate which was obtained by the LAAT algorithm is at the optimum in illumination and blur transformation. It also performs better in scale and rotation change. Moreover, LAAT has high accuracy in illumination, scale and rotation transformation conditions. The number of feature points extracted by LAAT is also close to other algorithms in all kinds of transformations. As the extraction time is not too long, it is expected to be used in real-time applications.

C. MATCHING RESULTS

The matching results are obtained by coarse matching and fine matching. We take the average of five test matching time as the final result.

From Fig. 8-9 and Fig. 11-12, it can be discerned that the number of matching pairs tends to decrease no matter the illumination decays or ambiguity increases. Fig. 8 and Fig. 11 depict the matches of LAAT are distributed more evenly than ORB and ORB+SURF when illumination changes. It tends to make better matches in brighter area

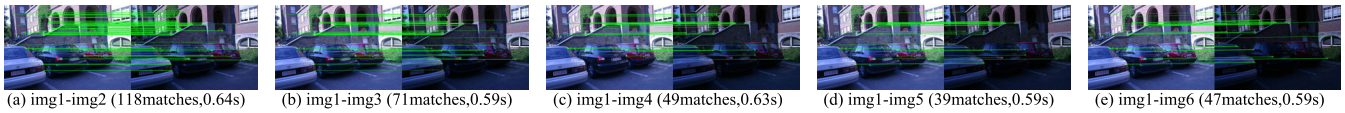


FIGURE 8. The matching results of the improved algorithm proposed in this paper under the change of illumination.

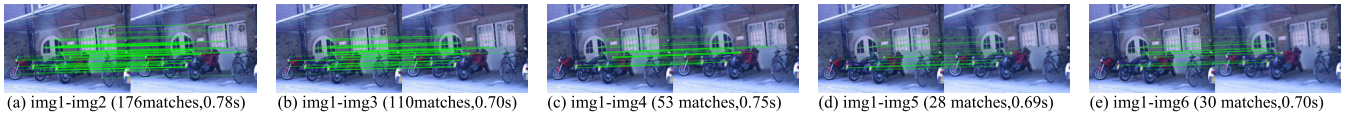


FIGURE 9. The matching results of the algorithm proposed in this paper under the change of blur.

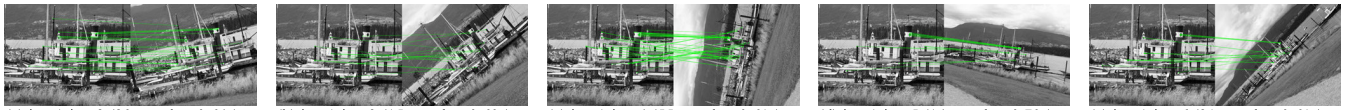


FIGURE 10. The matching results of the algorithm proposed in this paper under the change of zoom and rotation.

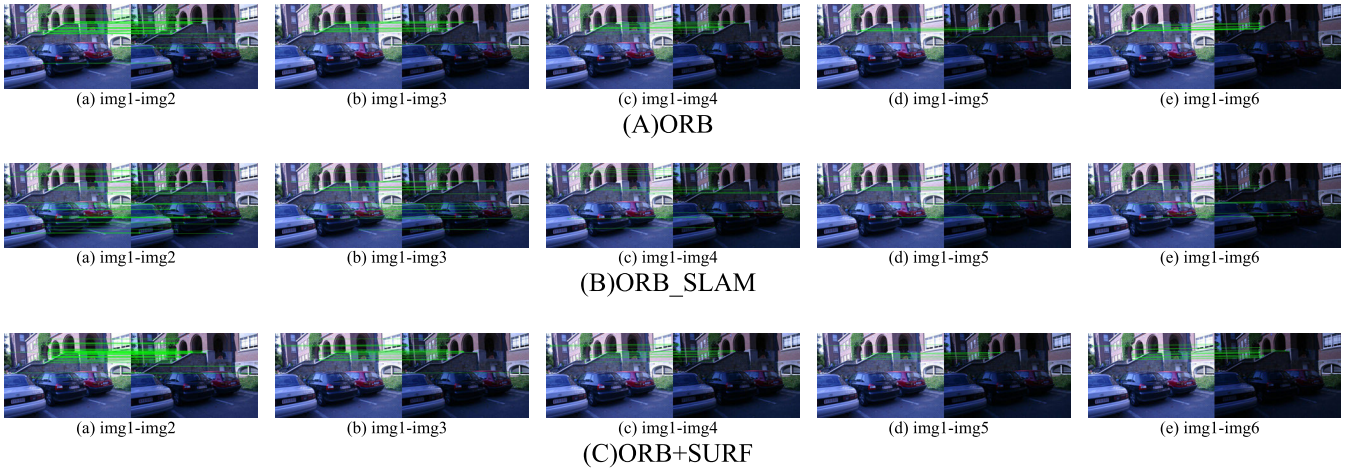


FIGURE 11. The matching results of the traditional ORB, ORB-SLAM and ORB+SURF under the change of illumination.

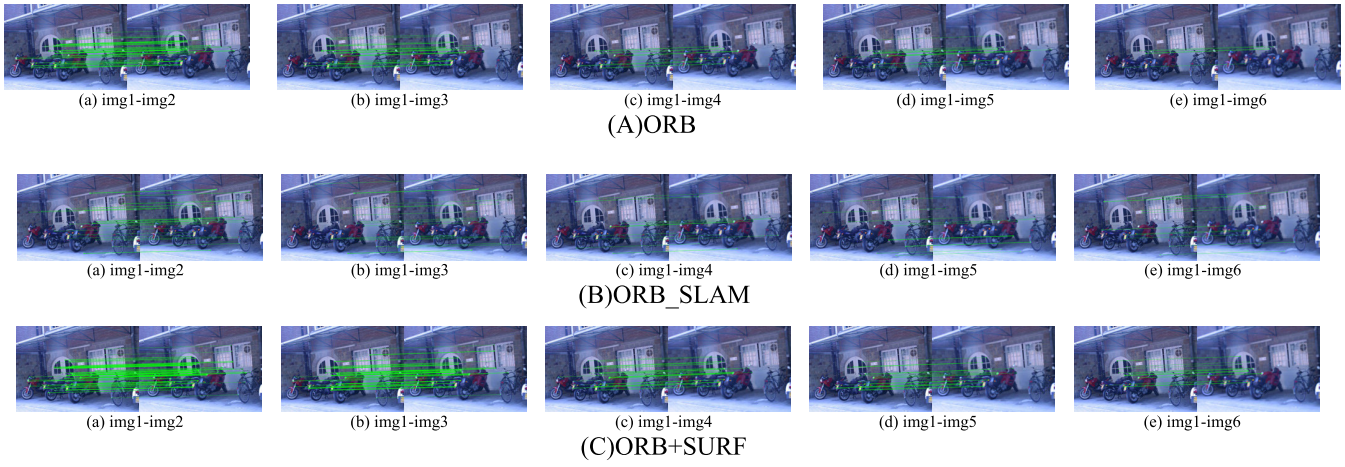


FIGURE 12. The matching results of the traditional ORB, ORB-SLAM and ORB+SURF under the change of blur.

rather than in dark area. As presented in Fig. 11(B), ORB-SLAM has a good performance in keeping matching pairs evenly distributed, either. However, its number of the matching pairs is too small. As shown in Fig. 10, Fig. 13 and Fig. 14(c), the matching results of all the algorithms seem sensitive to the change of zoom and rotation, and fluctuate

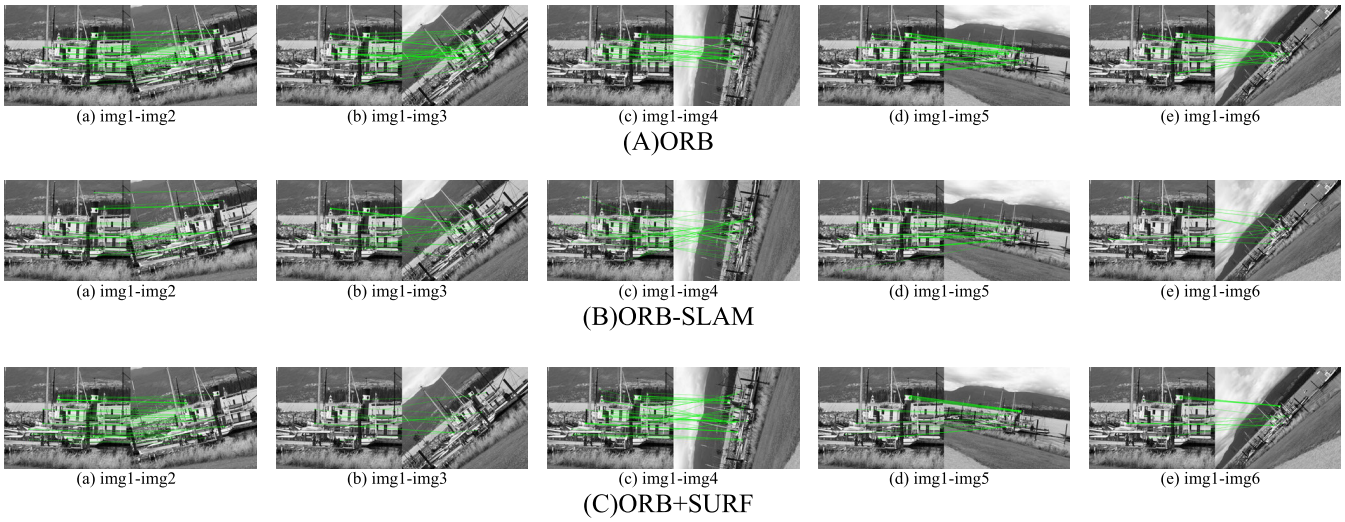


FIGURE 13. The matching results of the traditional ORB, ORB-SLAM and ORB+SURF under the change of zoom and rotation.

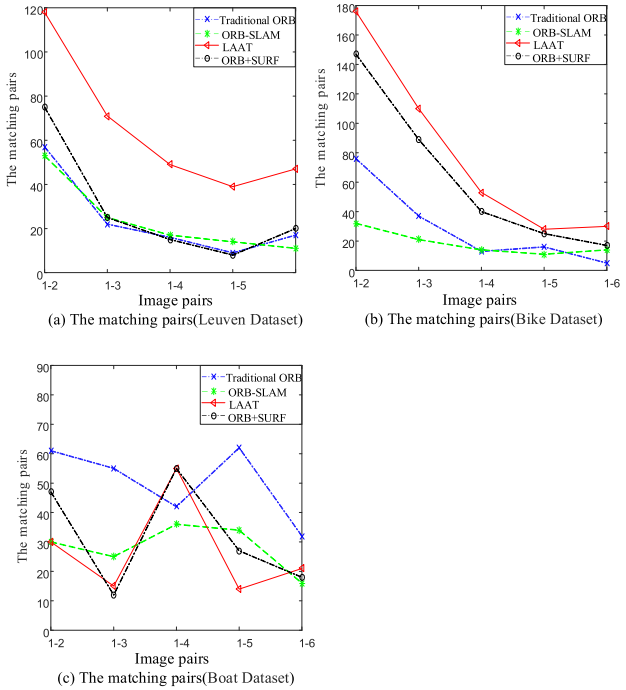


FIGURE 14. The number of the matching pairs by all the algorithms.

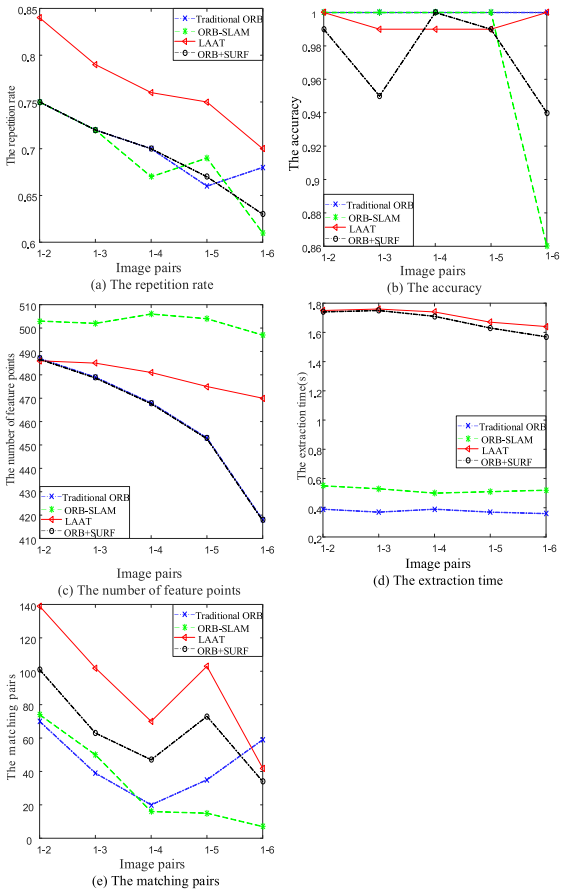


FIGURE 15. Extraction performance under varying brightness levels in the Leuven Dataset after reducing the image resolution.

widely. In these conditions, LAAT does not outperform the others. One interesting observation is that the numbers of the matching pairs of LAAT, ORB-SLAM and ORB+SURF reach maximum when the rotation degree is 90 degrees.

Fig. 14(a)(b) show that the numbers of the matching pairs of LAAT are significantly larger and are distributed more evenly than those of ORB, ORB-SLAM and ORB+SURF in the conditions of illumination and blur change. These characteristics make LAAT suitable for object tracking when the object moving between light and dark places.

D. THE INFLUENCE OF RESOLUTION

In order to inspect the influence of image resolution on the algorithm proposed in this paper, we reduce the image resolution of the above datasets and conduct relevant experiments.

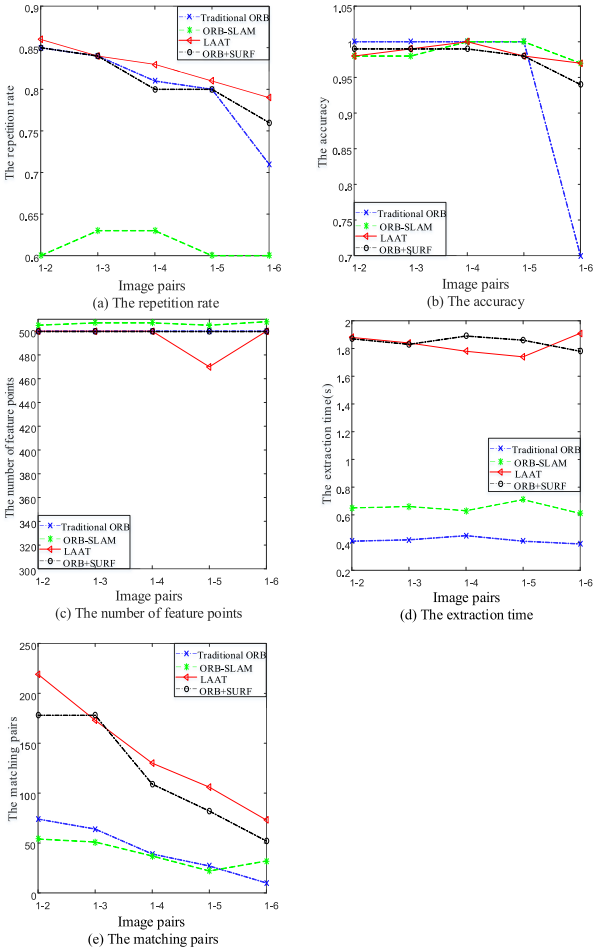


FIGURE 16. Extraction performance under varying blur levels in the Bike Dataset after reducing the image resolution.

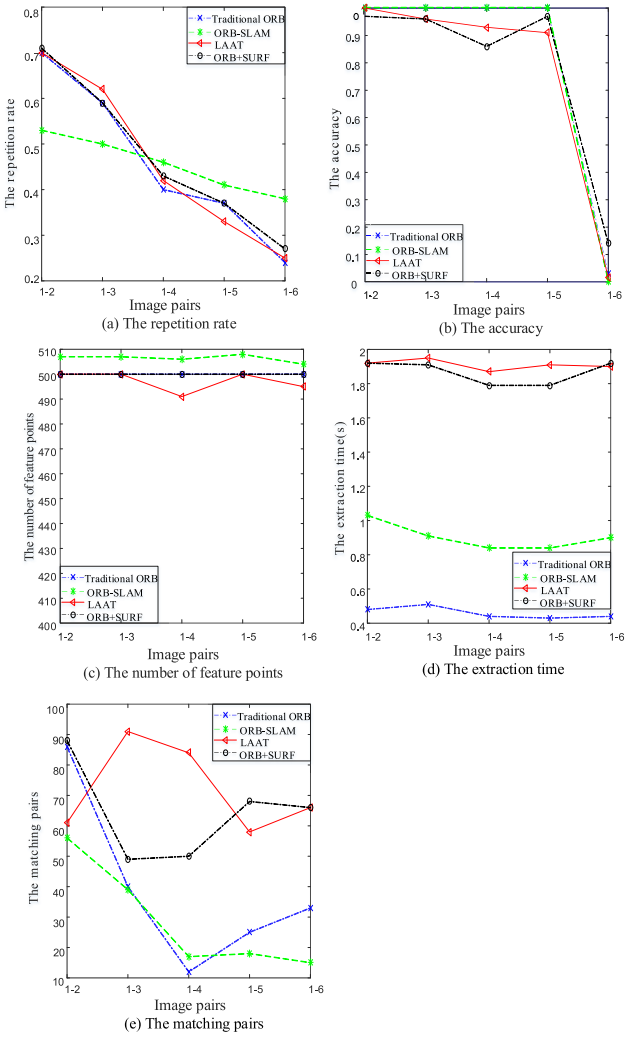


FIGURE 17. Extraction performance under varying rotation and scale levels in the Boat Dataset after reducing the image resolution.

1) ILLUMINATION CHANGE

As shown in Fig.15, Image resolution reduction can improve the repetition rates of the four algorithms in general. LAAT still shows its superiority compared with the other three algorithms in aspect of repetition rate It still has high accuracy rate over 0.99. LAAT outperforms the traditional ORB and ORB+SURF with respect to the number of feature points. The extraction time drops when the resolution is reduced. The number of matching pairs obtained by LAAT has increased and keeps higher than any of the other three algorithms.

2) BLUR CHANGE

From Fig.16, the repetition rate of LAAT keeps higher than others. The accuracy of LAAT shows more stabler in low resolution than in high resolution. For every algorithm, the number of feature points tends to be increased. The extraction time of LAAT drops to lower than 2 seconds. The number of matching pairs grows up and LAAT remains the highest.

3) ZOOM AND ROTATION CHANGE

As presented in Fig.17, when the image resolution is reduced, LAAT shows no obvious advantage regarding most of the above aspects. However, the number of matching pairs obtained by LAAT has been raised dramatically.

IV. CONCLUSION

Feature detection and matching are crucial aspects in many computer vision applications. They significantly affect the efficiency and accuracy of the applications. The experimental results have demonstrated that: (1)Compared with the algorithms, traditional ORB, ORB-SLAM and ORB+SURF, our proposed algorithm, LAAT, yields highest matching pair number in all datasets, and has higher repetition rate in image illumination change and blur change; (2)The number of matching pairs obtained by LAAT remains the highest in illumination change and blur change too; (3)When the image resolution is reduced, the robustness of illumination invariance and blur invariance for LAAT keeps the lead.

For the future work, we will focus on the improvement of the robustness for scaling and rotational changes and further

reduce process time to make it more suitable for real-time applications.

REFERENCES

[1] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd, "Real time localization and 3D reconstruction," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2006, pp. 363–370.

[2] D. Fleer and R. Möller, "Comparing holistic and feature-based visual methods for estimating the relative pose of mobile robots," *Robot. Auto. Syst.*, vol. 89, pp. 51–74, Mar. 2017.

[3] Z. Hu and Y. Jiang, "An improved ORB, gravity-ORB for target detection on mobile devices," in *Proc. 12th World Congr. Intell. Control Autom. (WCICA)*, Jun. 2016, pp. 1708–1713.

[4] E. Marchand, H. Uchiyama, and F. Spindler, "Pose estimation for augmented reality: A hands-on survey," *IEEE Trans. Vis. Comput. Graphics*, vol. 22, no. 12, pp. 2633–2651, Dec. 2016.

[5] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, "ORB-SLAM: A versatile and accurate monocular SLAM system," *IEEE Trans. Robot.*, vol. 31, no. 5, pp. 1147–1163, Oct. 2015.

[6] R. Mur-Artal and J. D. Tardos, "ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras," *IEEE Trans. Robot.*, vol. 33, no. 5, pp. 1255–1262, Oct. 2017.

[7] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1231–1237, Sep. 2013.

[8] S. Pillai and J. Leonard, "Monocular SLAM supported object recognition," 2015, *arXiv:1506.01732*. [Online]. Available: https://arxiv.org/abs/1506.01732

[9] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Sep. 1999, pp. 1150–1157.

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

[11] 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.

[12] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Proc. Eur. Conf. Comput. Vis.*, 2006, vol. 1, pp. 430–443.

[13] M. Cao, W. Jia, Y. Li, Z. Lv, L. Li, L. Zheng, and X. Liu, "Fast and robust local feature extraction for 3D reconstruction," *Comput. Electr. Eng.*, vol. 71, pp. 657–666, Oct. 2018.

[14] H. Kuang, X. Wang, X. Liu, X. Ma, and R. Li, "An improved Robot's localization and mapping method based on ORB-SLAM," in *Proc. 10th Int. Symp. Comput. Intell. Design (ISCID)*, Dec. 2017, pp. 400–403.

[15] F. Fausto, E. Cuevas, and A. Gonzales, "A new descriptor for image matching based on bionic principles," *Pattern Anal. Appl.*, vol. 20, no. 4, pp. 1245–1259, Nov. 2017.

[16] Z. Huang, Z. Wei, and G. Zhang, "BALG: An alternative for fast and robust feature matching," *J. Vis. Commun. Image Represent.*, vol. 60, pp. 129–139, Apr. 2019.

[17] S. Li, Z. Wang, and Q. Zhu, "A research of ORB feature matching algorithm based on fusion descriptor," in *Proc. IEEE 5th Inf. Technol. Mechatronics Eng. Conf. (ITOEC)*, Jun. 2020, pp. 417–420.

[18] R. Wang, Y. Shi, and W. Cao, "GA-SURF: A new speeded-up robust feature extraction algorithm for multispectral images based on geometric algebra," *Pattern Recognit. Lett.*, vol. 127, pp. 11–17, Nov. 2019.

[19] X. Wang, J. Zou, and D. Shi, "An improved ORB image feature matching algorithm based on SURF," in *Proc. 3rd Int. Conf. Robot. Autom. Eng. (ICRAE)*, Nov. 2018, pp. 218–222.

[20] A. Li, W. Jiang, W. Yuan, D. Dai, S. Zhang, and Z. Wei, "An improved FAST+SURF fast matching algorithm," *Procedia Comput. Sci.*, vol. 107, pp. 306–312, Jan. 2017.

[21] C. Cheng, X. Wang, and X. Li, "UAV image matching based on surf feature and harris corner algorithm," in *Proc. 4th Int. Conf. Smart Sustain. City (ICSSC)*, 2017, pp. 1–6.

[22] V. A, D. Hebbar, V. S. Shekhar, K. N. B. Murthy, and S. Natarajan, "Two novel detector-descriptor based approaches for face recognition using SIFT and SURF," *Procedia Comput. Sci.*, vol. 70, pp. 185–197, Jan. 2015.

[23] H. Li, H. Yang, and K. Chen, "Feature point extraction and tracking based on a local adaptive threshold," *IEEE Access*, vol. 8, pp. 44325–44334, 2020.

[24] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," in *Readings in Computer Vision*, 1987, pp. 726–740.

[25] F. Liu, Q. Lv, H. Lin, Y. Zhang, and K. Qi, "An image registration algorithm based on FREAK-FAST for visual SLAM," in *Proc. 35th Chin. Control Conf. (CCC)*, Jul. 2016, pp. 6222–6226.

[26] G. Majumder, M. K. Bhowmik, and D. Bhatacharjee, "Automatic eye detection using fast corner detector of north east indian (NEI) face images," *Procedia Technol.*, vol. 10, pp. 646–653, Jan. 2013.

[27] C. Jian, X. Xiang, and M. Zhang, "Mobile terminal gesture recognition based on improved FAST corner detection," *IET Image Process.*, vol. 13, no. 6, pp. 991–997, May 2019.

[28] H. Zhang, L. Xiao, and G. Xu, "A novel tracking method based on improved FAST corner detection and pyramid LK optical flow," in *Proc. Chin. Control Decis. Conf. (CCDC)*, Aug. 2020, pp. 1871–1876.

[29] P. L. Rosin, "Measuring corner properties," *Comput. Vis. Image Understand.*, vol. 73, no. 2, pp. 291–307, Feb. 1999.

[30] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[31] K. Mikolajczyk and C. Schmid, "An affine invariant interest point detector," in *Proc. 7th Eur. Conf. Comput. Vis. (ECCV)*, 2002, pp. 128–142.

ZHIMING CAI was born in 1977. He received the master's and Ph.D. degrees in communication and information system from Fuzhou University, Fuzhou, China, in 2004 and 2015, respectively. He is currently an Associate Professor with the School of Information Science and Engineering, Fujian University of Technology, China. His research interests include machine learning, computer vision, and application.

YIWEN OU received the B.S. degree from the Qingdao Institute of Technology, China, in 2018. She is currently pursuing the M.S. degree with the Fujian University of Technology, Fuzhou, China. Her research interests include stereo matching, object tracking, and binocular ranging.

YUFENG LING received the bachelor's degree in electrical engineering and automation from Qingdao University, Shandong, China. He is currently pursuing the master's degree in electrical engineering with the Fujian University of Technology, Fuzhou, China. His research interests include machine vision, 3D reconstruction, and plane detection based on deep learning.

JIAN DONG received the B.S. degree in electrical engineering and automation from the Hohai University Wentian College, Ma'anshan, China, in 2015. He is currently pursuing the M.S. degree in electrical engineering with the Fujian University of Technology, Fuzhou, China. His research interests include machine vision and image processing.

JIAN LU received the bachelor's degree from Jimei University, Xiamen, China, in 2019. He is currently pursuing the master's degree with the Fujian University of Technology, Fuzhou, China. His research interests include machine learning and path planning for robot.

HOWARD LEE received the M.E. degree in electrical and information engineering from The University of Sydney, Sydney, NSW, Australia, in 2000, and the Ph.D. degree in computer science from La Trobe University, Melbourne, VIC, Australia. From 2000 to 2001, he was a Research Assistant with the National Voice Centre and Signal and Multimedia Processing Laboratory, The University of Sydney. From 2003 to 2015, he was a Lecturer with Central Queensland University, Melbourne International Campus. His research interests include multimedia signal processing, pattern recognition, and network security. He was a member of the Association of Professional Engineers, Scientists and Manager Australia (APESMA). He was the Publication Chair of the IEEE-PCM 2000.