Improved ORB matching algorithm based on adaptive threshold

Shiji Li¹, Qing Wang* and Jiayue Li¹

¹Department of Computer Science and Technology, China Agricultural University, Beijing, 100083, China
* SY20193081482@cau.edu.cn

Abstract. Aiming at the problem of ORB feature matching algorithm extracting background pixels as feature points and matching wrong feature points in a complex background environment, an improved ORB algorithm based on adaptive threshold is proposed, and GMS algorithm is used to screen out mismatches in the feature matching stage. First, the algorithm calculates the mean and standard deviation of the image to be matched and the reference image. Then it inputs the obtained data into the adaptive threshold calculation stage to obtain the adaptive threshold. Finally it inputs the adaptive threshold into the feature extraction and matching stage. The experimental results show that the improved ORB algorithm reduces the number of features extracted from the background of the exhibits in the complex environment of the museum, and the matching algorithm combined with the ORB algorithm and GMS increases the correct matching on the basis of slightly shorter time than the original algorithm. The algorithm has strong robustness and real-time performance.

1. Introduction

Image feature matching is a very important step in augmented reality applications, and the result of feature matching is one of the keys to achieving object recognition. The image feature matching process is generally: extract feature points from the image; describe the feature points; finally perform image matching [1] to obtain the matching recognition result. Historically, researchers have proposed many methods for calculating image features. The SIFT algorithm [2] summarized and formally proposed by Canadian scholar David G. Lowe can accurately and stably extract feature points by constructing a 128-dimensional feature point description subset, but the computational complexity is high. The SURF algorithm [3] is improved on the basis of the SIFT algorithm, using Hessian matrix and Haar wavelet to extract the descriptor, reducing the dimension to 64 dimensions. The feature points extracted by these two algorithms have scale invariance and good matching performance, but the calculation time is long and the real-time performance is weak. Although many researchers have improved the SIFT algorithm and the SURF algorithm [4-5], the real-time performance is still relatively poor. Although the ORB algorithm [6] proposed in 2011 appropriately reduces the accuracy of feature points, it is two levels faster than SIFT in terms of speed and one level faster than SURF. Therefore, the ORB algorithm is a compromise algorithm in terms of feature point quality and speed. In recent years, many experts have improved the ORB algorithm from different perspectives. Literature [7] proposed an improved ORB algorithm based on the scale-invariant feature transform (SIFT) algorithm. Literature [8] proposed the SURB algorithm that combines ORB algorithm and SURF algorithm for feature point detection. Literature [9] proposed a method that uses ORB technology, and uses random sample consensus (RANSAC) as a post-processing step to eliminate redundant key points and noise, and improves ORB...
efficiency to provide a stable image recognition system.

The environment in which museum exhibits are located is generally an environment with dim background, complex lighting, and mirror reflections of visitors. Therefore, in order to identify the images of museum exhibits, the influence of these factors must be considered. Aiming at the factors affecting the identification of exhibits in the special environment of the museum, this paper proposes an improved ORB algorithm with adaptive threshold. The algorithm can automatically calculate the adaptive threshold according to the environment in which different exhibits are located, reducing the matching of background to feature points. In terms of feature matching, GMS (Grid-based Motion Statistics for Fast, Ultra-robust Feature Correspondence) [10] is used to match the feature points. The results show that the improved ORB algorithm can well remove the influence of background on feature point extraction and matching.

2. ORB Algorithm

The ORB algorithm is a feature point extraction and description algorithm with good real-time performance. It is composed of Oriented FAST key points and Steer BRIEF descriptors. Oriented FAST is an improved algorithm of FAST, which adds direction information to the extracted key point information. Rotated BRIEF is an improved BRIEF descriptor, which makes up for the shortcoming that the BRIEF descriptor has no rotation invariance.

2.1 Oriented FAST key points

The basic principle of the FAST algorithm is as follows: For any pixel p in the image, the FAST algorithm compares the pixel value I_p of this point with the pixel values of 16 points on the surrounding circle to determine whether the point is a feature point. There is a threshold t during the comparison process. When the pixel value of a point on the circle is greater than I_p+t, it means that the point is brighter than point p; when the pixel value of a point on the circle is smaller than I_p-t, it means that the point is darker than point p; when the pixel value of a point on the circle is between I_p-t and I_p+t, it means that the point is similar to point p. If there are consecutive n pixels on the circle that are brighter than p, or darker than p, then p is determined to be a feature point. n generally takes 12, and the commonly used values of n in other cases are 9 and 11. The feature point judgment formula is shown in formula (1) and formula (2), I_x is the gray value of any point on the circle, I_p is the gray value of the point to be measured, and N is the sum of the points on the circle conforming to the determination function. The schematic diagram of FAST algorithm feature point detection is shown in Figure 1.

\[
f = \begin{cases} 
1, & |I_x - I_p| > t \\
0, & \text{else}
\end{cases}
\]

\[
N = \sum_{x \in \text{around}(p)} f(I_x, I_p)
\]

In order to speed up the judgment process, I_p is generally only compared with the pixel values of 4 points at equal distances in Figure 1. Generally, the pixel values of 1, 5, 9, 13 in the cross direction are compared one by one. This method has been shown to have the same effect as comparing pixels from 16 surrounding points. If there is at least a pair of consecutive pixels whose brightness is higher or lower than p, then p is selected as the feature point.
FAST feature points are calculated by comparing the pixel differences between image points, so the speed is relatively fast, but this method also has some problems. First, the FAST feature points have no direction information. The ORB algorithm uses the gray-scale centroid method [11] to add a main direction to the feature points: The method first determines the center of mass of the local region of the feature points, and then constructs a direction vector with the center O of the local region pointing to the center of mass C. The moment defining a local area is

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

where $I(x,y)$ is the gray value at point $(x,y)$. The center of mass of the local region can be obtained by the moment:

$$C = \left( \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right)$$

where $m_{00}$ is the zero-order moment, and $m_{10}$ and $m_{01}$ are the first-order moments. The direction of the feature point can be defined as:

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

Secondly, the feature points of the FAST algorithm may not be feature points in some cases due to the visual distance problem. There is a problem of scale invariance. The ORB algorithm solves this problem by constructing an image pyramid and detecting key points on each layer of the pyramid.

### 2.2 Rotated BRIEF descriptor

Since the BRIEF descriptor does not have rotation invariance, the ORB algorithm uses the direction information of the calculated feature points to achieve rotation invariance. BRIEF is a binary coded descriptor whose description vector is defined as:

$$\tau(p;x,y) = \begin{cases} 1, & p(x) < p(y) \\ 0, & \text{else} \end{cases}$$

where $p(x)$ and $p(y)$ are the pixel values of the local region $p$ at the pixel points $x$ and $y$. Randomly select $n$ pairs of points $(x_i, y_i)$ to generate a binary string, and the generated feature descriptor can be expressed as

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

The ORB algorithm solves the problem that the BRIEF descriptor does not have rotation invariance is: randomly select $n$ point pairs to form a matrix $Y$, as shown below

$$Y = \begin{bmatrix} x_1, \ldots, x_n \\ y_1, \ldots, y_n \end{bmatrix}$$

A new description matrix can be obtained by rotating the affine transformation matrix $R$, which is determined by the main direction of the feature points detected:

$$S_\theta = R_\theta S = \begin{bmatrix} \cos \theta \sin \theta \\ -\sin \theta \cos \theta \end{bmatrix} \begin{bmatrix} x_1, \ldots, x_n \\ y_1, \ldots, y_n \end{bmatrix}$$

### 3. GMS Algorithm

The basic idea of GMS is to distinguish between correct matches and incorrect matches by counting the total number of matches in the neighborhood of each pair of matching points. Extract $N$ and $M$ feature points from the two images $I_a$ and $I_b$ respectively, and the matching probability of each point is independent, and the probability of correct matching is set to $t$. Through brute force matching, the matching set $X = \{x_1, x_2, \ldots, x_N\}$ is obtained. Suppose that the matching $x_i$ contains $n$ and $m$ feature points in the fields $a$ and $b$ in the two images, and $f_i$ is a feature point in $a$. If $x_i$ is a correct match, the probability that the matching point of $f_i$ falls in the region $b$ is
\[ p_t = p(f_a^{ab}) = p(f_a^{ab}) + p(f_a^{bf}) + p(f_a^{fb}) = t + (1-t) \frac{\beta m}{M} \]  

In the formula: \( T^{ab} \) means that \( x_i \) is a correct match, \( f_a^b \) means that the matching point of \( f_a \) falls in the \( b \) area, \( f_a^t \) means that \( f_a \) is a correct match, \( f_a^{fa} \) means that \( f_a \) is incorrectly matched, and \( \beta \) is an adjustment factor. Similarly, if \( x_i \) is a false match, the probability that the match point of \( f_a \) falls in the region \( b \) is

\[ p_f = p(f_a^{fb}) = p(f_a^{fb}) + p(f_a^{bf}) = \frac{\beta (1-t)m}{M} \]  

The total number of matching points in the \( X_i \) field is its support degree, which is set as \( S_i \), and the probability distribution of \( S_i \) is a binomial distribution:

\[ S_i \sim B(n, p_t), \quad \text{a and b are the same area} \]

\[ S_i \sim B(n, p_f), \quad \text{a and b are different areas} \]

In order to improve calculation efficiency, GMS generally divides the image into 20×20 grids in the actual calculation process, and calculates the support of each grid. The statistical area of the support of each grid is the grid and the 8 grids around it:

\[ S_{ij} = \sum_{k=1}^{9} |x_{ikjk}| \]  

In the formula: \( |x_{ikjk}| \) represents the number of matching points of the two grids in the two pictures, and \( S_{ij} \) also conforms to the binomial distribution.

The above formula shows that each grid has support in its neighborhood, but the support distribution is different, and its distribution is bimodal. So by choosing a suitable threshold, you can effectively judge whether the grid is acceptable. Set the threshold as:

\[ T = \alpha (n_{ij})^{1/2} \]  

In the formula: \( n_{ij} \) is the average of the number of matching points in the 9 grids, \( \alpha \) is the parameter for adjusting the threshold, and Bian Jiawang et al. set \( \alpha \) to 6 in the paper [10].

### 4. Improved ORB algorithm with adaptive threshold

Aiming at the problem that the ORB algorithm is easily affected by the illumination and background in the museum environment, an improved algorithm of adaptive threshold is proposed. According to the characteristics of the exhibit images, this algorithm adaptively selects the threshold \( t \) for the two images for feature point matching. Through analysis, the size of the image standard deviation reflects the difference in contrast. For the two images for feature matching, the following method of automatically selecting the threshold is introduced:

\[ t = k \sigma + b \]  

where \( t \) is the FAST adaptive threshold, \( \sigma \) is the absolute value of the deviation between the standard deviations of the two images, \( k \) and \( b \) are the adaptive parameters, and the value is determined according to experimental data. Through a large number of experimental statistics, the value of \( b \) is related to \( k \) and the absolute value of the deviation between the mean values of the two images, as shown in the following formula:

\[ b = \begin{cases} 0, & \mu_1 < k \\ k, & k \leq \mu_1 < 2k \\ 2k, & 2k \leq \mu_1 \leq \mu_2 \\ |b|, & \mu_1 > \mu_2 \end{cases} \]  

where \( \mu \) is the absolute value of the deviation between the mean values of the two images, \( \mu_1 \) is the mean...
value of the exhibit image to be matched, and $\mu_2$ is the mean value of the reference exhibit image.

5. Analysis of results

In order to verify the performance of the algorithm in this paper, two experiments are carried out: ORB-GMS image feature matching comparison experiment and adaptive threshold algorithm comparison experiment. The experimental simulation environment is JetBrains PyCharm Edu 2019.1, and the computer system is Windows 10 Professional Edition [Intel(R) Core(TM) i7-7700K CPU @ 4.20GHz, 8G memory]. The experimental test data adopts actual exhibit images taken at the National Museum. The experimental images are all taken under the perspective of simulating visitors’ normal viewing of the exhibits, with different points of view, shadow, background, etc.

5.1 ORB-GMS image feature matching comparison experiment

In order to verify the performance of image feature point matching after the ORB algorithm is combined with GMS, the feature points obtained by the original ORB algorithm are matched using Hamming distance, and then the brute force matching method selected by the knn algorithm is compared with the ORB-GMS algorithm. The experiment uses two algorithms to perform feature point matching process with a set of matching pictures. The experimental results are shown in the following figure. The number of feature points matching and time consumption are shown in the table below.

![Figure 2. The first set of ORB brute force matching results](image1)

![Figure 3. The first set of ORB-GMS matching results](image2)

![Figure 4. The second set of ORB brute force matching results](image3)

![Figure 5. The second set of ORB-GMS matching results](image4)

**Table 1. Comparison of matching results**

| Algorithm           | Image   | Number of matching points | Time (ms) |
|---------------------|---------|---------------------------|-----------|
| ORB Brute-Force Match | Figure 1 | 908                       | 289.2     |
| ORB-GMS             | Figure 2 | 2905                      | 285.2     |
| ORB Brute-Force Match | Figure 3 | 738                       | 290.2     |
| ORB-GMS             | Figure 4 | 3331                      | 288.2     |

According to the data in the image and the table, the ORB-GMS algorithm has improved the number of feature point matching and the number of correct matches, and the entire algorithm takes about 1.3% less time than the traditional ORB brute force matching method. The time is almost the same as the ORB
brute force matching method. In some ORB-GMS experimental results, the matching results of the feature points are not very correct. This is due to the influence of the environmental factors of the exhibits. Based on the experimental results, an adaptive threshold algorithm experiment was subsequently carried out.

5.2 Adaptive threshold algorithm comparison experiment

In order to verify the feature point extraction effect of the adaptive threshold algorithm, ORB-GMS algorithm was used to extract and match feature points in each group of images without setting thresholds or setting adaptive thresholds. The experimental results are shown in the figure below.

![Figure 6. Match results without an adaptive threshold](image1)
![Figure 7. Matching result with adaptive threshold of 40](image2)

6. Conclusion

Aiming at the problem that the ORB algorithm is susceptible to the influence of background, shadow and other factors in the extraction of feature points of museum exhibits, resulting in a large number of non-exhibit feature points, this paper proposes an improved ORB algorithm with adaptive threshold. The algorithm calculates an adaptive threshold based on the contrast difference of different images, eliminates many unnecessary feature points, and saves time and cost. At the same time, in terms of image feature point matching, this paper combines ORB algorithm with GMS to screen the preliminary matching results. The experimental results show that the improved algorithm improves the matching accuracy and time in the museum exhibit environment than the traditional ORB algorithm, and has stronger robustness and real-time performance. But for exhibits with inconspicuous features or too dense features, the effect of this algorithm has not reached the ideal effect in the experiment, which is also a goal that needs further research in the future.

Acknowledgments

Thanks to my mentor Wang Qing. Thanks for their valuable suggestions for my article.

References

[1] Gao J, Wu Y, Wu K, et al. (2012) Image matching method based on corner detection. Chinese Journal of Scientific Instrument, 34(8):1717-1722
[2] Lowe D G. (2004) Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision(S0920-5691),60(2):91-100.
[3] Herbert B, Andreas E, Tinne T, et al. (2008) Speeded-Up Robust Features (SURF). Computer Vision and Image Understanding(S1077-3142), 110(3): 346-359.
[4] Di N, Li G, CHEN C. (2014) Image matching technology research based on normalized DOG features. Journal of Electronic Measurement and Instrumentation,28(6):585-560.
[5] Lu M, Zhao F, Li N. (2015) Feature extraction of high-resolution SAR images based on bag-of-word model. Foreign Electronic Measurement Technology,34(6):62-69.
[6] Rublee E, Rabaud V, Konolige K, et al. (2011) ORB: An efficient alternative to SIFT or SURF. In: International Conference on Computer Vision. Barcelona. 2011(2):2564-2571.
[7] Qin Y, Xu H, Chen H. (2014) Image feature points matching via improved ORB. In: 2014 IEEE
International Conference on Progress in Informatics and Computing. New York. 2014: 204-208.

[8] Dai X, Lang L, Chen M. (2016) Research of image feature point matching based on improved ORB algorithm. Journal of Electronic Measurement and Instrumentation, 30(2): 233-240.

[9] Vinay A., Rao AS, Shekhar VS, et al. (2015) Feature extraction using ORB-RANSAC for face recognition. Procedia Computer Science, 70: 174-184.

[10] Bian J, Lin W., Matsushita Y, et al. (2017) GMS: grid-based motion statistics for fast, ultra-robust feature correspondence. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Hawaii. 2828-2837.

[11] Rosin P. (1999) Measuring Corner Properties. Computer Vision and Image Understanding, 73(2): 291-307