An augmented reality image registration method based on improved ORB

Mingzhi Cheng, Luyue Zhang, Long Liu*
Beijing Institute of Graphic Communication, Beijing, China

Corresponding Author Email: liu88204277@163.com

Abstract. In the process of Augmented Reality (AR) image registration, the traditional ORB (oriented FAST and rotated BRIEF) algorithm has low registration rate and poor real-time performance. In this paper, an improved AR image registration method based on improved ORB is proposed. Firstly, the calibration image and video frame image feature points are obtained by the improved FAST feature detection algorithm. Then, the binary descriptor of BRISK, which using the custom domain sampling pattern is used for feature description, and the scale invariance of the traditional ORB algorithm is improved. Finally, the random sampling consistency (RANSAC) algorithm is used to eliminate the wrong matching point pairs and optimize the feature matching. Experiments show that compared with the AR image registration method described by the traditional ORB algorithm and the FREAK feature, the registration rate of the proposed algorithm is increased by 1.1% and 8.4%, and the generation time is reduced by 0.13s and 0.12s, respectively. The experimental results show that the AR image registration method proposed in this paper can obtain higher feature point registration rate, and has better real-time performance, which can better meet the application needs of AR image registration.

1. Introduction
AR is an emerging technology that enables users to interact with virtual content seamlessly in a real-world environment, and has been widely used in different fields. Currently, the technical points of AR mainly include: three-dimensional (3D) registration, Human Computer Interaction (HCI), and Virtual Reality (VR) fusion display. The key to the realization of VR fusion display technology is how to complete image registration by combining computer vision. Domestic and foreign scholars generally divide the problem into two categories: based on markers and no marker. The former uses a simple detection algorithm to quickly recognize black and white, non-overlapping square images. Although some practical problems are solved, it is difficult to connect with real-world objects. This has largely limited its application in the AR field. However, the no marker registration method compensates for the deficiencies of the marker-based method, at the same time, it exhibits greater inclusiveness to the shape and texture of the recognition target. Therefore, the image registration method based on no marker has been paid much attention by scholars, and the improvement of its algorithm tends to be mature.

SIFT and SURF algorithm as image registration algorithms which based on no marker proposed earlier than others and have relatively excellent performance, but there is great time consumption in the meantime. Later, many researchers proposed PCA-SIFT, R-ASURF and other improved algorithms, which improved the performance of detection algorithms to a certain extent. On the one hand, Edward Rosten et al. [1] proposed FAST feature point detection algorithm, which greatly
improved the time consumption. Michael Calonder et al. [2] proposed BRIEF binary descriptor to work out the problems of memory resources and time consumption. On the other hand, with the in-depth research and development of feature detection algorithm, ORB [3] algorithm that combined with improved FAST and BRIEF emerged at the right moment and became a classic algorithm in the field of feature detection.

Although ORB algorithm has excellent performance in robustness and effectiveness, it still has some defects with poor scale invariance and high error rate. In this paper, the ORB algorithm was improved to reduce the error rate of the traditional ORB algorithm, and random sampling consistency (RANSAC) algorithm was used to optimize the feature matching after the initial matching of feature points was completed, and the outer points were removed, further improving the image registration rate.

The structure of this paper is organized as follows: Section 2 shows the improvement and application of ORB algorithm, as well as some other algorithms could applied in AR. In section 3 obtains the principle and flow of ORB algorithm, while section 4 explains the improved ORB image registration method and the matching optimization method based on RANSAC. Finally, the analysis of experimental results and conclusions are discussed in section 5 and section 6.

2. Related work

2.1. Improvements of ORB algorithm
ORB algorithm and its improvement algorithm are commonly used in the field of computer vision. Jiang et al. [4] put forward an improved algorithm for posture estimation of automatic battery replacement system. This computing framework combined with ORB and EPnP algorithm that could be applied to automatic assembly of computer vision, with good accuracy and reliability. But it still needs to be resolved that how to make the number of key points of detection objects more appropriately meet the needs. Zhong et al. [5] proposed a visual tracking algorithm combining ORB feature points and target color models to solve the problem of CAMShift algorithm tracking failure in color similar background. Zhang et al. [6] were able to continue tracking and drawing when SLAM system tracking failure, and the number of maps keyframes was higher than ORB-SLAM. Mur-Artal et al. [7] proposed ORB-SLAM2 solution with lightweight positioning mode and excellent precision.

In addition, ORB algorithm has been widely used in the fields of image stitching [8], 3D reconstruction [9], medical imaging [10] and so on. It also plays an important role in AR. Ren et al. [11] had put forward the implementation method of AR system based on ORB algorithm, which demonstrates the usefulness of ORB algorithm in AR real-time system. Wang et al. [12] combined with ORB and LINE-MOD algorithms to construct descriptors and update camera posture, and then implemented an AR assembly auxiliary system based on unmarked tracking methods. However, the feature descriptors of this method did not solve the problem of scale invariance.

2.2. Algorithm application in AR
In the application of AR, researchers often select appropriate algorithms and apply them in AR projects to improve performance so that satisfied the needs of practical applications as much as possible. Soyoung et al. [13] introduced SURF algorithm in the maintenance phase of the construction project, which is used to link 3D objects into the maintenance work and improve the information utilization rate of 3D objects. Zhang et al. [14] proposed an image recognition strategy based on SIFT algorithm and applied it to mobile AR games. Their study has achieved better response speed and information expansion than traditional AR system. Moreover, it could meet the high requirements of efficiency and accuracy in AR system. Wu et al. [15] studied about the AR alignment method in surgical image-guided surgery. As result of introducing stochastic perturbation technique into traditional ICP (Iterative Closest Point) alignment method, they reduce the alignment error range to less than 3 mm. Takacs et al. [16] use a new descriptor trained by machine learning method to provide EDD descriptor for AR applications, which had the characteristics of good adaptability to unstable
environmental conditions and was satisfied with most environmental changes. With the rapid development of AR technology, the application of image registration algorithm in the field of AR also has a broad prospect.

To sum up, the research of traditional ORB algorithm and its improved algorithm provides better image registration solution for AR technology. Therefore, the ORB algorithm is selected as the research object in this paper, and the practical problems in AR application are realized by improving the ORB algorithm.

3. Traditional ORB algorithm

3.1. Feature detection

The ORB algorithm is improved on the basis of FAST feature detection algorithm. The FAST algorithm detects the image feature point by comparing the light and dark information of the center pixel and the surrounding pixels. However, the characteristic points obtained by FAST detection do not perform well in terms of scale characteristics and rotation invariance. In view of this, in order to obtain certain scale characteristics, the ORB algorithm selects FAST detectors to detect feature points at each level of the image pyramid, and then sorts the feature points according to the size of Harris response values. The first N point is selected based on the threshold. The corner point response function can be expressed as equation (1):

\[ R = \text{Det}(M) - k \times \text{trace}(M) \times \text{trace}(M) \]  

(1)

In order to obtain the rotational invariant feature, the first moment is used to calculate the local direction through the weighted average of the pixel size in the local region. The moment definition of image block in the literature is given as equation (2):

\[ m_{p,q} = \sum_{x,y} x^p y^q I(x,y) \]  

(2)

Where, \((x,y)\) is located in the circular region with radius \(r\), and the point \((x,y)\) has gray value \(I(x,y)\). The centroid coordinates of the image block can be expressed as equation (3):

\[ C = (C_x, C_y) = \left( \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \]  

(3)

By connecting the center point of the corner point of the image block with the centroid of the gray value, the vector formed from the feature point to centroid. The direction of the feature point is defined as equation (4):

\[ \theta = \arctan \left( \frac{m_{01}}{m_{10}} \right) = \arctan \left( \frac{\sum_{x,y} y I(x,y)}{\sum_{x,y} x I(x,y)} \right) \]  

(4)

3.2. Feature description

The basic idea of BRIEF feature description can be described as equation (5):

\[ \tau(p; x, y) = \begin{cases} 1, & p(x) < p(y) \\ 0, & p(x) \geq p(y) \end{cases} \]  

(5)

Randomly select \(n\) pixel pairs in the neighborhood of a feature point, and calculate the gray value of all point pairs according to the binary rule. Then generate binary string of length \(n\), which is the feature descriptor of the feature point. Once the ORB feature detection algorithm is derived, the
feature descriptor can be calculated by using the feature point with scale characteristics. The feature
description of ORB algorithm is based on the feature description of BRIEF algorithm, which gives the
rotation invariance by Steered BRIEF method, and trains the test results through the greedy search
method, then uses the statistical learning method to re-select the point-to-point set. As a result, the
ORB algorithm has achieved better diversity and lower correlation. In addition, the 31*31 pixels
neighborhood of the feature point selects 5*5 pixels sub-window to compare the pixel block pairs by
the integral plot. The feature description of the feature points obtained by the above steps reduces the
sensitivity to noise.

The specific steps of the ORB's characterization algorithm are:
1. Define the image neighborhood that has been processed by Gaussian distribution as \( p \).
2. Establish \( 300K \) feature points test set and consider 31*31 pixels neighborhoods for each point
in it.
3. The value of the pair is represented by the average of the grayscale of the 5*5 sub-window of
feature points, and all the pairs are taken in each point's 31*31 neighborhood, then the binary string
of 300K feature points is obtained to form matrix \( Q \).
4. Find the average \( s \) of each binary string, and resort the matrix vectors in the order of \( |s - 0.5| \) to
form a vector \( T \).
5. Use the greedy search method to get the description:
   (1) Put the first amount of \( T \) into \( R \) and remove it from vector \( T \);
   (2) Compare the correlation between the next quantity in \( T \) and all column vectors in \( R \), and
deposit the low correlation into \( R \);
   (3) Repeat the previous step until the vector number of \( R \) reaches 256, and the rBRIEF feature
description is constructed.

3.3. Feature matching
The feature matching of the ORB algorithm uses the Hamming distance as the evaluation element, and
the Hamming distance is the smallest pair. Through the binary feature description operator constructed
in the section 4.2, the feature descriptors of a group of images are randomly selected as equation (6).
The feature matching of the ORB algorithm is determined by the similarity of descriptors. The XOR
operation of two descriptors is performed by the Hamming distance, and the result of the summation
represents the similarity which can be recorded as equation (7).

\[
K_2 = x_0 x_1 \cdots x_n, K_2 = y_0 y_1 \cdots y_n \tag{6}
\]

\[
D(K_1, K_2) = \sum_{i=0}^{n} x_i \oplus y_i \tag{7}
\]

The smaller the \( D(K_1, K_2) \) is, the higher the similarity of the feature point pairs.

4. Improved image registration method
4.1. Image registration algorithm based on improved ORB
The ORB algorithm benefits from the FAST feature detector and binary descriptor, and has the
advantages of fast operation speed and small resource space, but it still has the disadvantages of high
error rate and low registration. The feature descriptor of ORB algorithm does not solve the problem of
scale invariance, which greatly affects the application of image registration algorithm in AR scenes. In
this paper, an image registration method based on improved ORB is proposed. The feature index of the
image to be registered is detected by the oFAST (oriented FAST) feature detection algorithm. The
scale invariance is provided by the BRISK feature descriptor for the ORB algorithm, and the
robustness is constructed. The binary descriptor with scale invariance not only solves the defect of
ORB feature descriptor in scale invariance, but also obtains higher registration rate, so that it has better real-time performance.

Firstly, the feature description of improved ORB algorithm constructs the feature descriptors, which using a custom domain sampling mode. The sampling model consists of concentric circles. Concentric circles with different radii are created around the feature points, and then take a certain number of equal-spacing pixel sampling points on each circle, as shown in figure 1. In order to eliminate the image aliasing interference caused by this sampling mode, Gaussian smoothing is used at each sampling point, and the radius of red dotted line corresponds to the Gaussian kernel standard deviation of the smoothed sampling point brightness.

![BRISK sampling model](image)

Figure 1. BRISK sampling model.

The two-two combination of \( N \) sampling points have \( N(N - 1)/2 \) point pairs, and all combinations form a sampling point set \( A \) of equation (8). The sampled pixel pair \( (p_i, p_j) \) is Gaussian filtered to obtain the pixel gray value \( I(p_j, \sigma_j) \) and \( g(p_i, p_j) \), the local gradient value of the sampled feature point is described as equation (9):

\[
g(p_i, p_j) = (p_j - p_i) \frac{I(p_j, \sigma_j) - I(p_i, \sigma_j)}{||p_i - p_j||^2}
\]  

(8)

In addition, in the calculation of the feature point direction, the long direction point pair is used to confirm the main direction. According to the distance between the pair of sampling points, the short distance point pair subset \( S \) and the long distances point pair subset \( L \) are defined as follows in equation (9) and equation (10):

\[
S = \{(p_i, p_j) \in A ||p_i - p_j|| < \sigma_s \} \subseteq A
\]

(9)

\[
L = \{(p_i, p_j) \in A ||p_i - p_j|| > \sigma_L \} \subseteq A
\]

(10)

Where, \( t \) is defined as the scale at which the feature point is located, \( \sigma_s \) is a short-distance threshold, usually equal to 9.75\( t \), \( \sigma_L \) is a long-distance threshold, usually equal to 13.67\( t \). The main direction of the feature point can use the long distances point to the subset to be calculated by equation (11). Finally, when the algorithm constructs the feature descriptor, the sampling region is firstly rotated according to the main direction of the feature point, and the rotation angle is determined by equation (12) as follows:

\[
g = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \frac{1}{l_{(p_i, p_j) \in L}} \sum_{(p_i, p_j) \in L} g(p_i, p_j)
\]

(11)
\[ \theta = \tan^{-1}(gy, gx) \]  

(12)

Then, the grayscale value of the sample point pair is compared in a subset \( S \) of the short-distance point pair, resulting in a binary feature descriptor of 512 bits based on equation (13) as follows:

\[
b = \begin{cases} 
1, & I(p_i^\theta, \sigma_i) > I(p_j^\theta, \sigma_j) \\
0, & \text{others} 
\end{cases}
\]  

(13)

Therefore, the image registration algorithm based on the improved ORB utilizes the BRISK construction descriptor, the circular sampling structure of this descriptor brings rotation invariance and scale invariance for the selection calculation of sampling points. Compared to the traditional ORB feature description of random sampling does not fully consider the scale characteristics, the improved ORB algorithm can make the feature descriptor contain more scale and rotation characteristics.

4.2. Matching optimization method based on RANSAC

Feature matching is performed after feature detection and feature description steps. Common feature point matching methods includes: Brute-force based matching and Flann based matching. The first matching method is also called brute matching. This method selects the key points by selecting the image, then test each key point of the detected image, followed by returning to the nearest key point. In the result of the second fast nearest neighbor matching algorithm only performs well in the large data set, the first matching method is used in this paper.

In this paper, the Bruce-force-based feature matching method is used to calculate the Hamming distance of the binary descriptor, and the nearest distance of the obtained distance is used as a matching point, then the RANSAC algorithm is adopted for the feature points obtained through violent matching to exclude the exclusion points, to achieve the purpose of matching optimization. The random sampling consistency algorithm uses an iterative approach from a set of data sets, matches the given mathematical model parameters, and estimates the feature points and the error rate of the model to evaluate the accuracy of the model [17]. In this paper, the matching efficiency of feature points is optimized by matching based on random sampling consistency, and the matching images after eliminating the errors are obtained. The experimental effect is shown in figure 2. It can be seen from the comparison of figure 2(a) and figure 2(b) that the image registration method with RANSAC algorithm eliminates most of the wrong point pairs and achieves the purpose of matching optimization.

Figure 2. (a) Image registration effect without matching optimization
5. Experimental results and analysis

This experiment runs on the system of windows 10, VS 2010, OpenCV 2.4.3 and OpenGL are used to implement image registration methods based on improved ORB and RANSAC. In the experiment, the hardware configuration included Intel Core i7-7700K CPU 4.20GHz, 16G memory, and Logitech 720p camera.

In this paper, a group of images in figure 2 is registered. When the threshold of feature point extracted is 1000, the experimental results are shown in Table 1. It can be seen from Table 1 the improved algorithm is compared with the image registration algorithm described by ORB feature and the image registration algorithm described by FREAK feature. Obviously, the improved algorithm can obtain more matching points and higher registration accuracy. Also, the generation is short in time and has good real-time performance. Overall, the algorithm in this paper can better meet the needs of augmented reality image registration.

As we can see in the Table 1, when the threshold of feature points is set to 200, 500 and 1000, the experimental results of the above cases are given in figure 3.

| Applying algorithms  | Matching points | Registration rate/% | Consuming time/s |
|----------------------|-----------------|---------------------|------------------|
| ORB                  | 600             | 60                  | 1.75             |
| FREAK                | 493             | 52.7                | 1.74             |
| Improved algorithm   | 600             | 61.6                | 1.62             |
Respectively, there are the number of feature point matching, the registration rate of feature points, and the consumption time of generated registration. The above graph shows that within the three threshold ranges set, the number of feature point matching obtained by the improved algorithm in this paper is approximately compared with the other two traditional algorithms, which can satisfy the experimental needs excellently. When the improved algorithm is compared with the AR image registration algorithm that described by the ORB feature, the registration rate is increased by 1.1%, and the generation time is reduced by 0.13s. Compared with the AR image registration algorithm using FREAK feature description, the registration rate is improved by 8.4% and the generation time is reduced by 0.12s. The experimental results show that the proposed algorithm outperforms the traditional algorithm in the above aspects and it is certain that this improved algorithm is a better augmented reality image registration method. It provides a good foundation for the subsequent implementation of augmented reality scenarios.

Finally, this paper applies the improved image registration method based on ORB algorithm to realize AR application of mechanical drawing, and provides teaching aids for mechanical drawing teaching and learning. It has obtained good application effect, and the rendering effect of AR scene is shown in figure 4.

6. Conclusions
Aiming at the defects of low registration rate and high error rate of traditional ORB algorithm, this paper proposes an improved AR image registration method based on ORB algorithm and RANSAC. Firstly, the improved FAST feature detection algorithm is used to detect the feature points, and then the improved feature description of ORB algorithm is used to obtain the binary feature descriptor.
After the feature initial matching phase is completed, we optimize the matching based on the algorithm of RANSAC. Finally, an AR system framework is constructed by using the improved algorithm proposed in this paper, and the AR applications can be implemented more efficiently. Experimental results show that the proposed algorithm outperforms the traditional algorithm and has good real-time performance. However, the method of this paper has a good effect on the experiment of viewing angle change image, and it is not stable to other environmental changes such as light intensity. In the future work, we will further optimize the algorithm of this paper to satisfy the actual requirements of various environments. In addition, research on camera pose calculation will be continued to further enhance the scene’s effect.

**Acknowledgments**
This work was supported in part by the New Engineering Project of the Ministry of Education in 2018: "Exploration and practice of engineering talent training mode integrating the three disciplines of 'engineering and art' under the background of Internet +" under Grants 22150118004/001, the Key Teaching Reform Project of Beijing Institute of Graphic Communication in 2017: "Research on interdisciplinary talent cultivation mode of new media" under Grants 22150117092, and the Lingchuang Class Project of Beijing Institute of Graphic Communication in 2019 under Grants 22150119018.

**References**

[1] Rosten, E., Porter, R., Drummond, T. (2008) Faster and better: a machine learning approach to corner detection. IEEE Trans Pattern Anal Mach Intell, 32(1), 105-119.

[2] Calonder, M., Lepetit, V., Strecha, C., Fua, P. (2010) Brief: Binary robust independent elementary features. Proc of the 11th European Conference on Computer Vision.

[3] Rublee, E., Rabaud, V., Konolige, K., Bradski, G. (2011) ORB: an efficient alternative to SIFT or SURF. Proc of IEEE International Conference on Computer Vision.

[4] Jiang, J.B., Wu, F., Zhang, P.F., Yang, Y.Y. (2019) Pose Estimation of Automatic Battery-Replacement System Based on ORB and Improved Keypoints Matching Method. Applied science, 9(2): 237.

[5] Zhong, H.M., Wang, W., Zhang, H.H. (2015) Visual tracking algorithm combined with ORB features and color models. Pattern Recognition and Artificial Intelligence, 28(01): 90-96.

[6] Zhang, J.H., Wang, Y.Y., Wang, Z.Y., Chen, S.Y., Guan, Q. (2018) Single-eye simultaneous positioning and map recovery fusion technology in the building. China Journal of Image and Graphics, 23(03): 372-383.

[7] Raul, M.A., Juan, D.T. (2017) ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. IEEE Transactions on Robotics, 33(5): 1255-1262.

[8] Yuan, R.F., Liu, M., Hui, M., Zhao, Y.J., Dong, L.Q., Kong, L.Q., Cai, Z. (2018) Real-time video stitching based on ORB features and SVM. Proceedings of SPIE, 10752.

[9] Shao, A.J., Qian, W.X., Gu, G.H., Li, C., Mao, C. (2014) Research on image registration algorithm based on computationally integrated imaging. Infrared technology, 37(05):398-403.

[10] He, X.J., Ling, Y.G., Zhang, Y.X., Liang, Z.S. (2018) Multimodal image registration simulation under digital medical imaging technology. Computer simulation, 35(12): 166-170.

[11] Ren, J., Zhou, Y., Yu, Y., Du, S.D., Wang, Z.Q. (2012) AR real-time system implementation based on ORB's natural characteristics. Computer Application Research, 29(09):3594-3596.
[12] Wang, Y. Zhang, S.S. Yang, S. He, W.P., Bai, X.L. (2016) Mechanical assembly assistance using marker-less augmented reality system. ASSEMBLY AUTOMATION, 38(1): 77-87.

[13] Moon, S.Y., Yun, S.Y., Kim, H.S., Kang, L.S. (2015) Improved Method for Increasing Maintenance Efficiency of Construction Structure Using Augmented Reality by Marker-Less Method. Journal of the Korean Society of Civil Engineers, 35(4): 961-968.

[14] Zhang, B.P. (2017) Design of mobile augmented reality game based on image recognition. Image Video Process, 1.

[15] Wu, M.I., Chien, J.C., Wu, C.T., Lee, J.D. (2018) An Augmented Reality System Using Improved-Iterative Closest Point Algorithm for On-Patient Medical Image Visualization. SENSORS, 18(8):2505.

[16] Takacs, A., Manuel, T.A., Jesus, C.P.O., Edgar, A.R.A. (2018) Dedicated feature descriptor for outdoor augmented reality detection. PATTERN ANALYSIS AND APPLICATIONS, 21(2): 351-362.

[17] Cacciari, P.P., Futai, M.M. (2017) Modeling a shallow rock tunnel using terrestrial laser scanning and discrete fracture networks. Rock Mechanics and Rock Engineering, 50(5): 1217-1242.