

RESEARCH ARTICLE

# A novel image registration approach via combining local features and geometric invariants

Yan Lu, Kun Gao\*, Tinghua Zhang, Tingfa Xu

Key Lab of Photoelectronic Imaging Technology and System, Ministry of Education of China, Beijing Institute of Technology, Beijing, China

\* [gaokun@bit.edu.cn](mailto:gaokun@bit.edu.cn)



**OPEN ACCESS**

**Citation:** Lu Y, Gao K, Zhang T, Xu T (2018) A novel image registration approach via combining local features and geometric invariants. PLoS ONE 13(1): e0190383. <https://doi.org/10.1371/journal.pone.0190383>

**Editor:** Dalin Tang, Worcester Polytechnic Institute, UNITED STATES

**Received:** April 6, 2017

**Accepted:** December 13, 2017

**Published:** January 2, 2018

**Copyright:** © 2018 Lu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was funded by the Natural Science Foundation of Beijing Municipality (grant No. 4152045, <http://www.nsf.gov.cn/publish/portal1/>) to KG, National Natural Science Foundation of China (grant No. 61527802, <http://www.bjnsf.org/>) to TX, and the National High Technology Research and Development Program of China (grant No. 2014AA7026082, <http://program.most.gov.cn/>) to TX. The funders had no

## Abstract

Image registration is widely used in many fields, but the adaptability of the existing methods is limited. This work proposes a novel image registration method with high precision for various complex applications. In this framework, the registration problem is divided into two stages. First, we detect and describe scale-invariant feature points using modified computer vision-oriented fast and rotated brief (ORB) algorithm, and a simple method to increase the performance of feature points matching is proposed. Second, we develop a new local constraint of rough selection according to the feature distances. Evidence shows that the existing matching techniques based on image features are insufficient for the images with sparse image details. Then, we propose a novel matching algorithm via geometric constraints, and establish local feature descriptions based on geometric invariances for the selected feature points. Subsequently, a new price function is constructed to evaluate the similarities between points and obtain exact matching pairs. Finally, we employ the progressive sample consensus method to remove wrong matches and calculate the space transform parameters. Experimental results on various complex image datasets verify that the proposed method is more robust and significantly reduces the rate of false matches while retaining more high-quality feature points.

## Introduction

Image registration is mainly for obtaining the transformation parameters between images taken at different times from different angles and sensors with translation, rotation, scaling and/or distortion to get the best match in the pixel layer [1]. Image registration is a fundamental issue for many computer vision technologies, such as image restoration, targeting, tracking, image stitching, image fusion, 3D reconstruction, and pattern recognition [2,3]. It is an important preliminary step to improve the accuracy and validity of the above problems.

Image registration techniques play an increasingly important role in various fields. For military applications, to improve precision strike capabilities, various images from different sensors, such as infrared, radar, and hyperspectral imaging, require high-precision registration [4,5]. The requirements are similar for civil use equipment for security monitoring, traffic

role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

control, wide field imaging and panoramic imaging [6,7]. In agriculture, image registration is also critical. For example, three-dimensional imaging devices, such as photonic mixer detectors, are used to capture image depth information. Those images are then matched with images from color cameras to more accurately record vegetation growth conditions [8]. Image registration is also an urgent issue for remote sensing applications, such as the investigation of geological disasters and the exploration of complex terrain. Image registration is often applied to optical images, synthetic aperture radar (SAR) images, and other multispectral images to achieve more surface feature information. Image registration has essential research value, especially in the medical field. Image registration technology is often used to obtain the exact location of organs and tissues from different image sources, such as X-ray computed tomography (CT), magnetic resonance imaging (MRI), single photon emission computed tomography (SPECT), and positron emission tomography (PET) [9,10], which enables reliable diagnosis and treatment. Therefore, research on image registration has important theoretical significance and practical application value.

The grayscale distribution, image details, and noise interference are all uncertain factors that challenge image registration. For instance, as in Fig 1A, many algorithms will fail to match the images with starry backgrounds since the images have a high dynamic range and lack texture features. The exposure time and the signal-to-noise ratio are also different. They all will lead to a high rate of false matches and even match failure. When the images have a large grayscale contrast and considerable noise interference, the mismatch rate will increase rapidly, as shown in Fig 1B. Therefore, the anti-interference ability and robustness of the algorithms are important factors in image registration. Moreover, accurate image registration is still challenged by images that have many similar details but small and detailed differences between the current and target images (see Fig 1C).

Considering the problems above, this paper proposes a fast and robust image registration method, as shown in Fig 2. We utilize an improved feature detection and description method by modifying the computer vision-oriented fast and rotated brief (ORB) algorithms to provide scale invariance. Then, this approach is combined with the distribution of key points to form a robust method of feature selection and matching. For feature selection, a new distance window is set after considering the feature point location distribution and distance constraints. Then, high-quality feature points are extracted following the Hamming distance criterion and bidirectional matching constraint from the K nearest neighbor (KNN). We establish the feature description vector based on the geometric invariance, which differs from traditional matching methods. A cost function is constructed to evaluate the similarity of vectors in different dimensions and obtain better matched point pairs. Thus, even if there are relatively few feature points in images, we can still accurately register the images. Finally, space transformation parameters for image pairs are calculated after using the progressive sample consensus (PROSAC) algorithm to eliminate error matching; then, the final feature point matching relationship can be obtained.

## Related works

Previous researchers have proposed many excellent image registration methods, which can be classified as area- or feature-based. Regardless of the approach, the aim is to obtain the invariants from image pairs and find matching relations. Ideally, these invariants should not be affected by light, noise, or geometric deformation [11]. In recent years, methods based on local invariant feature points have become the major focus due to their superior performance.

Since the concept of corner detection was proposed by Moravec [12], many detection operators have been investigated, such as Harris, Shi-Tomasi, SUSAN, and others [13]. Lowe first

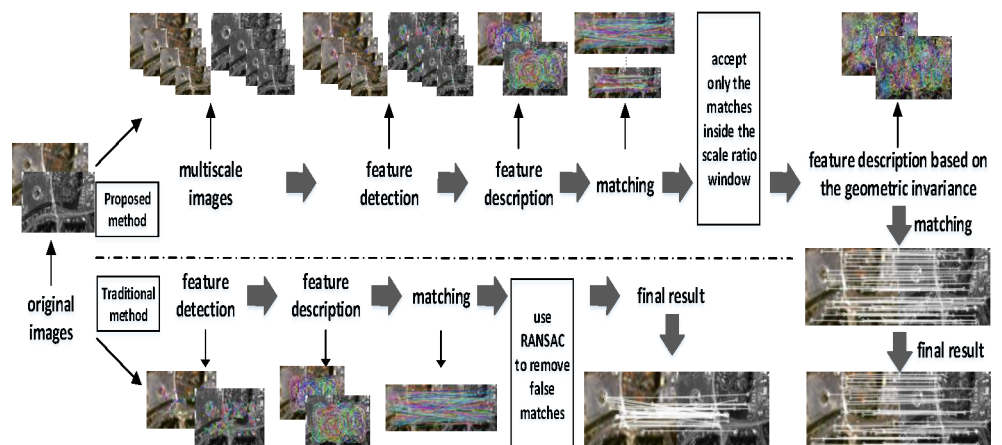


**Fig 1. The mismatches in image registration.**

<https://doi.org/10.1371/journal.pone.0190383.g001>

proposed the scale invariant feature transformation (SIFT) algorithm [14]. The SIFT algorithm provided a huge improvement in accuracy for corner detection. It is robust to light changes, noise, and affine transformation and can achieve sub-pixel accuracy [15]. Bay et al. [16] improved SIFT's efficiency and proposed the speed up robust feature (SURF) algorithm, which used Hessian matrices and distributed descriptors. SURF allows multiple images in a scale space to be processed simultaneously and does not require image subsampling. Therefore, it can effectively reduce descriptor dimensionality and significantly improve calculation speed while guaranteeing accuracy.

To apply the feature point detection algorithm in real time, Rosten and Drummond [17] proposed a new method called the features from accelerated segment test (FAST) algorithm. FAST compares surrounding pixels to obtain key points using machine learning. Therefore, it is simple, effective and easily ported to embedded systems [18]. Calonder et al. [19] proposed the BRIEF descriptor by comparing the PCA, LDA and other feature dimensional reduction methods. It reduces the time needed to generate feature descriptors by calculating and matching binary strings. Subsequently, FAST and BRIEF were redesigned by Rublee et al. They proposed the ORB algorithm at the 2011 IEEE International Conference on Computer Vision, which provided significant advantages in performance and speed [20].



**Fig 2. Proposed image registration framework.**

<https://doi.org/10.1371/journal.pone.0190383.g002>

In recent years, the registration technology based on image features has many applications. Zhang et al. [21] registered medical images by establishing a key feature model to describe the features and matching the corresponding points via a geometric constraint. Because the traditional methods were insufficient to achieve adequate results under different image deformations, Kahaki et al. [22] proposed an invariant feature matching method to overcome the limitations by measuring the dissimilarity of the features through the path based on eigenvector properties. Then, they achieved the registration of high resolution IKONOS satellite images. Li et al. [23] proposed an approach to robustly build key point mappings on multispectral images, and a similarity transformation was considered to account for the misalignment between two images. Lee et al. [24] proposed an application of the SIFT algorithm to stitch cervical-thoracic-lumbar (C-T-L) spine magnetic resonance (MR) images, and the results indicated that it can be improve diagnosis capabilities.

The local feature detection and matching methods described above have good real-time performance, noise immunity, robustness, and other positive characteristics, but image registration remains a challenging research topic [25]. Registration accuracy, reliability and computational time are three important characteristics that constrain universal registration methods in different circumstances. The traditional methods, such as SIFT or SURF, have excellent performance and high precision. However, when the images lack texture features, the feature points are difficult to extract and describe, and the matching result will be similar to Fig 1A. For images with rich textures, although it is possible to extract a large number of high-quality feature points, key point selection and precise matching still have many problems, particularly for images that are captured at different times, phases, or using different sensors, such as medical images. These images often have relative distortion, deformation, and/or uneven illumination. Traditional matching methods based on global or local features are limited by key point quantity and quality, which makes it difficult to guarantee precise results. In addition, it is difficult to describe the invariance of an image with many similar feature points by traditional methods. Therefore, we propose solving these problems by combining the modified local features and geometric invariants.

## Feature detection and description

Although many modified registration methods based on image features can theoretically enhance computational efficiency, ORB always performs better in complex scenarios [6]. It is fast, effective and accurate. Therefore, during feature detection and description, we improve the ORB algorithm to extract higher quality feature points to satisfy more complex applications.

In this step, we use the improved FAST-9 algorithm to detect features. A feature discriminant response function  $T$  is defined as

$$T = \sum_{i \in \text{circle}(p)} |G(i) - G(p)| > \xi \quad (0 < \xi < 255), \tag{1}$$

where  $G(p)$  and  $G(i)$  are the grayscale values at  $p$  and its neighboring points and  $\xi$  is the threshold value. Here, we set the threshold to 40. When comparing the 16 neighbor points, if there are 9 consecutive points in the circular boundary of the neighborhood and their grayscale values are larger than  $\xi$ , it is judged to be a feature point. Then, the corner response function proposed by Harris was used to select from the identified feature points.

Then, the directions of feature points need to be calculated. For any of the feature points, the neighbor moments  $M_{pq}$  of the neighborhood pixels and the centroid of these moments  $C$

can be expressed as

$$M_{pq} = \sum_{x,y} x^p y^q I(x,y) \text{ and } C = \left( \frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right), \tag{2}$$

where  $x, y$  are the positions of feature points and the centroid can be calculated by  $M_{pq}$ . The angle between the feature point and centroid is set as the dominant orientation, expressed as  $\alpha = \text{atan2}(M_{01}, M_{10})$ , where  $\text{atan2}$  is the quadrant-aware version of  $\arctan$ . These feature points provide directional invariance but are not scale invariant. This will be improved later.

After detecting the feature points, the improved BRIEF descriptor is used to describe these features. First, the point pairs are randomly generated in image patches. Let  $(p(m), p(n))$  be the grayscale values for a point pair  $p$ , where each point pair corresponds to a binary string test  $\lambda$ .

$$\lambda(p; m, n) := \begin{cases} 1 & \text{if } p(m) < p(n) \\ 0 & \text{otherwise} \end{cases}. \tag{3}$$

Then,  $k$  point pairs are randomly selected for generating a binary string. The feature descriptors  $D_k$  are expressed as

$$D_k(p) := \sum_{1 \leq i \leq k} 2^{i-1} \lambda(p; m_i, n_i). \tag{4}$$

These descriptors are based on the pixel values and are easily affected by noise. Therefore, a neighboring sub-window of feature points is defined, and the pixel value is replaced by comparing the gray-level integration of the sub-window. To ensure the descriptor has rotational invariance,  $n$  pairs of features are chosen at points  $(x_i, y_i)$  and form a matrix  $S$ .

$$S = \begin{pmatrix} x_1 & x_2 & \dots & x_i & \dots & x_n \\ y_1 & y_2 & \dots & y_i & \dots & y_n \end{pmatrix} \quad (i = 1, 2, \dots, n), \tag{5}$$

where  $(x_i, y_i)$  are the coordinates of the points. Then, using the dominant orientation and affine transformation matrix obtained in the feature detection stage, a new feature description matrix  $S'$  and descriptor  $D'$  can be calculated by rotating the affine transformation matrix such that

$$S' = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \cdot S, \tag{6}$$

$$D'(p, \alpha) := D_k(p) | (x_i, y_i) \in S'. \tag{7}$$

Finally, using greedy search, 256 pixel pairs with minimum correlation can be found to describe the features.

To achieve scale invariance, traditional methods use the multiscale partitioning before feature detection, and then feature detection and extraction are separately performed. However, this often results in considerable mismatched features between low- and high-resolution images, which reduces the final matching rate. The SIFT algorithm also suffers from the same problem. Bastanlar et al. [26] proved this issue and proposed a preprocessing SIFT (PP-SIFT) solution.

In this work, we proposed an optimized method to reduce mismatches. The following processing steps are added to detect and match features of multiscale images.

Step 1: For images at a high resolution, we adopt a Gaussian low-pass filter and down sampling both horizontally and vertically.



**Fig 3. The matching under the proposed improved method.**

<https://doi.org/10.1371/journal.pone.0190383.g003>

Step 2: Apply ORB matching to images and plot the histogram of scale ratios.

Step 3: Form the histogram of scale differences and define a window  $|H_{\max} \pm \omega|$  around the peak of histogram  $H_{\max}$ . Parameter  $\omega$  is set between 0.20 and 0.35. The matches with scale differences outside this window are rejected.

Step 4: Extract the correct scale ratio ( $d$ ) from the histogram as the mean of the most dominant Gaussian in the mixture.

Step 5: Accept only the matches with a scale ratio between  $0.6d$  and  $1.4d$ .

Fig 1B is the matching result of the original ORB algorithm, and Fig 3 is the result of our improved method. Both methods remove false matches using the random sample consensus (RANSAC) [27,28]. The result shows that mismatches are significantly reduced, and a large number of correct matches remain after the optimization.

### Feature selection and matching method

The feature descriptor mentioned above is a binary string that provides increased storage and matching speed. The traditional matching method often utilizes the brute force (BF) [29] algorithm to match the feature points, which is followed by the RANSAC algorithm to eliminate mismatches. It is effective for normal scenes [30]. However, for complex applications, such as the matching of medical images or remotely sensed images, the generation of a large number of interference points can frequently occur and lead to a high mismatch rate [22]. Furthermore, when the feature points are sparse, it is difficult to guarantee an accurate matching [31]. To address this challenge, we propose new constraints of rough selection according to the distribution of feature points. We establish a new feature description vector and matching criterion based on geometrical relationships and employ the PROSAC algorithm for accurate matching.

### Constraint by feature distribution

Let  $hd_1$  and  $hd_2$  be the binary strings of feature descriptors for two images constructed by the ORB algorithm.

$$\begin{cases} hd_1 = p_1 p_2 p_3 \cdots p_{255} \\ hd_2 = q_1 q_2 q_3 \cdots q_{255} \end{cases}, \quad (8)$$

where  $p$  and  $q$  are the descriptors of two images. Then, the Hamming distances  $D$  of the image

features is the XOR operation for the descriptors

$$D(hd_1, hd_2) = \sum_{i=0}^{255} p_i \oplus q_i \tag{9}$$

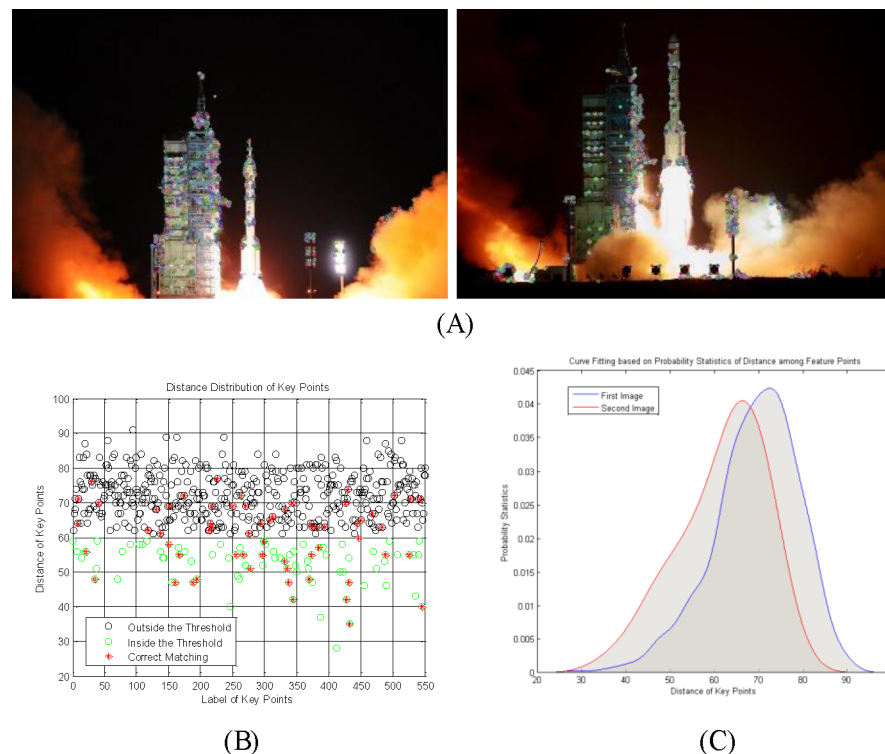
In traditional methods, feature points with Hamming distances smaller than a previously set threshold [32]  $\epsilon$  are

$$Match_j = D_j \ (D_j < \epsilon), \tag{10}$$

where  $D_j$  are the distances of the  $j$ th match point pairs and  $Match_j$  are the selected matches. This method is applicable for images with relatively even distributions of features. However, if the image includes an energy-focused region or a region with an intensive feature distribution, it is difficult to define an appropriate threshold to avoid mismatches between nearby feature points. Since these points below the threshold are very similar, the image contains many low-quality feature points, especially in areas with strong noise interference.

Fig 4 is the distribution of feature points for two images. Fig 4A shows the detected feature points. The selection result is expressed by green circles when the threshold satisfies  $\epsilon = 60$ , and the red stars (\*) show the reliable matches, as shown in Fig 4B. The traditional method filters most of the low-quality feature points, but many correct matchings are also removed. The threshold parameter is unstable and unreliable as a selection standard.

Fig 4C shows the fitted curves based on the probability statistics of the feature point distances. It has a large overlap area around the mean with a corresponding high contact ratio, and most of the reliable feature points are distributed in the overlap area. Therefore, the mean



**Fig 4. The distribution of feature points.** (A) input images and the feature points; (B) the distance distribution of feature points; (C) curve fitting based on probability statistics of the distances.

<https://doi.org/10.1371/journal.pone.0190383.g004>

centered constraint condition is set as a rough selection in this paper. We define the mean of root mean distance  $\bar{D}$  as

$$\bar{D} = \sqrt{\frac{1}{n} \sum_j^n D_j}. \tag{11}$$

The selection window  $R$  is centered on  $\bar{D}$  and defines the matching points to be retained.

$$\bar{D} - \varepsilon_1 \leq R \leq \bar{D} + \varepsilon_2 \quad (\varepsilon_1 > 0, \varepsilon_2 > 0), \tag{12}$$

where  $\varepsilon_1$  and  $\varepsilon_2$  are the upper and lower limits of  $R$ , respectively, and can be modified according to the image feature distribution density. Here, we set  $\varepsilon_1$  to the medium value between the minimum distance and  $\bar{D}$  and set  $\varepsilon_2$  to the medium value between the maximum distance and  $\bar{D}$ .

The distance constraint can remove significant errors and retain most of the reliable feature points, but it still needs further screening. KNN bilateral matching is employed in the following steps to select more reliable matching points.

Let  $p_i$  be a key point in the current frame, and let  $p_{j1}$  and  $p_{j2}$  be the two nearest matches of Hamming distance in the corresponding reference frame. Their distance are  $D(p_i, p_{j1})$  and  $D(p_i, p_{j2})$ , which are optimal and sub-optimal, respectively. Similarly, there are two corresponding matching points for a key point in the reference frame with distances  $D(p_j, p_{i1})$  and  $D(p_j, p_{i2})$ . There are also two candidate matching points based on the descriptor distance in another image.

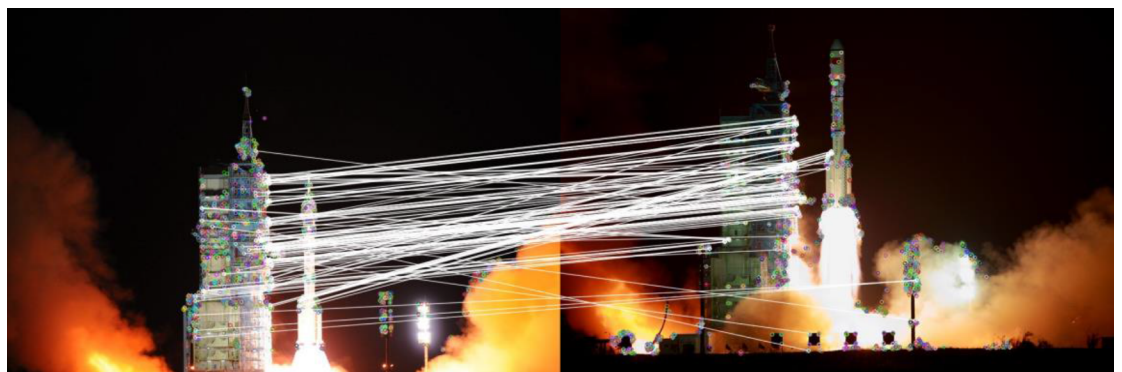
The ratio of the optimal and sub-optimal value is used as the selection condition for the two images' feature points. Two better quality sets of key points can be obtained using

$$R_1' = \frac{D(p_i, p_{j1})}{D(p_i, p_{j2})} < t, \text{ and} \tag{13}$$

$$R_2' = \frac{D(p_j, p_{i1})}{D(p_j, p_{i2})} < t. \tag{14}$$

Here, we set  $t$  to 0.65 according to the experiments. Finally, matches that simultaneously satisfy both conditions are the respective optimal matches.

After applying the distance constraint of the selection window and bilateral matching, many false matches are filtered without using RANSAC, as shown in Fig 5.



**Fig 5. The matching result after the distance constraints.**

<https://doi.org/10.1371/journal.pone.0190383.g005>



## A matching method based on geometric invariants

After the rough selection, the remaining feature points are more robust with higher quality. However, many mismatches may still exist for complicated situations. The main reason is that not all of the points identified by the Hamming distance match are correct matching pairs, and key point distance is only one factor considered for matching. The method has some limitations, especially when lacking texture and details or when there are few feature points [33]. It is also difficult to exactly match image pairs that have large deformations.

Therefore, this paper considers geometric invariance as a reliable matching factor, and a new method based on geometric invariants is proposed to provide further selection and matching of feature points. The geometric invariance was often used to describe the shape of objects, such as the shape context algorithm proposed by Belongie et al. [34]. It considers the object's shape and contour in the image and makes full use of contextual information for image sequences. The algorithm uses the log polar histogram to describe the contour sampling distribution, and it is widely used for digital recognition, trademarks, and the like. Here, we briefly introduce the principles of this method and then propose our method.

Let the set  $\mathbf{P} = \{p_1, p_2, \dots, p_m\}$  with  $m$  sampling points describe an object's shape. The log polar histogram  $H_i(m)$  of the other  $m-1$  sampling points is calculated as the shape contextual descriptor for each point.

$$H_i(m) = \#\{p_j \neq p_i : (p_j - p_i) \in \text{bin}(m)\}, \tag{15}$$

and the log polar transformation (LPT) can be expressed as

$$\rho = \log \sqrt{(x - x_0)^2 + (y - y_0)^2}, \tag{16}$$

$$\alpha = \arctan\left(\frac{y - y_0}{x - x_0}\right). \tag{17}$$

Then, divide  $\rho$  into five equal parts, divide  $\alpha$  into twelve equal parts, and form  $k$  sections. Every point has its own distribution relative to the others, so the number of sampling points in each sub-sector domain can be used as the similarity criterion. Accordingly, using the matching cost function  $F_{i,j}$  between feature points  $p_i$  and  $p_j$ , we get

$$F_{i,j} = \frac{1}{2} \sum_{m=1}^k \frac{[H_i(m) - H_j(m)]^2}{H_i(m) + H_j(m)}. \tag{18}$$

The feature point matching problem can be converted to match the weighted undirected bipartite graph. Finally, using the Hungary algorithm, we can find the optimal match and minimum cost value. Thus, the key points can be easily matched.

The feature of shape context can be easily extracted. The image scale and rotation transformation in the Cartesian coordinate system can be converted into the translation in the log coordinate system using LPT. Therefore, it has good scale and rotational invariance.

This algorithm has advantages for matching object shape, but the result may be affected by image noise and edge detection. Additionally, when objects are deformed, the matching accuracy and stability may be compromised [35]. The algorithm also requires that a point set must be the subset of a larger group, which is difficult to satisfy for image registration. Therefore, this paper utilizes the underlying theory of this method and proposes a new model to filter and match the feature points.

We define the matching point set obtained after coarse selection as  $\mathbf{R} = \{p_1, p_2, \dots, p_n\}$ . We calculate the distance from each key point  $p_i$  to the other  $n - 1$  key points  $p_j$  without dividing

the sharp histogram. The distances  $D_i(n)$  between these points can be expressed as

$$D_i(n) = \|p_i - p_j\|_2, p_j \in \mathbf{R}, i \neq j. \tag{19}$$

Assuming that  $k$  is the number of feature points, a feature description matrix with  $k \times (k - 1)$  dimensions can be obtained.

If the dimension is different between two images, the similarity among feature vectors cannot be measured by Eq (18). Therefore, an improved descriptive model is proposed. Let  $m$  and  $n$  be the number of feature points in the two images, with the feature vectors

$$\begin{cases} D_i = (D_{i1}, D_{i2}, \dots, D_{im-1}) \\ D_j = (D_{j1}, D_{j2}, \dots, D_{jn-1}) \end{cases} \tag{20}$$

Then, a new matching cost function  $F_{i,j}$  that considers the feature point distances is defined as

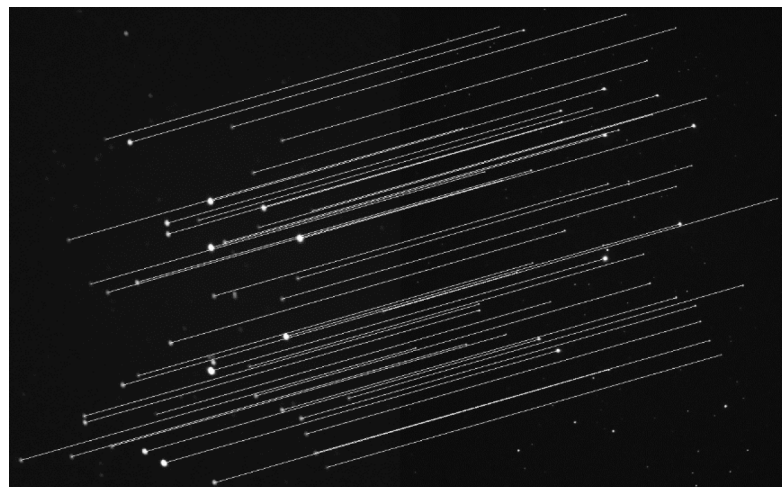
$$F_{i,j} = \sum_{s=1}^{m-1} \sum_{t=1}^{n-1} \exp\left(\frac{[D_i(s) - D_j(t)]^2}{\sigma^2}\right), \tag{21}$$

where  $D_i(s)$  and  $D_j(t)$  are the feature vectors of the current image and target image, respectively.  $\sigma$  is the controllable distance error threshold set it to 1 in this paper.

Finally, we construct the binary search trees for every key point according to the cost function. Then, we calculate the ratio of the previous  $K$  nodes and compare them with threshold  $T$  to judge whether they are an acceptable matching point pair.

$$R_m = \frac{F_{i,j}(max)}{F_{i,j}(m)} > T, \quad (m = 1, 2, \dots, K), \tag{22}$$

where  $F_{i,j}(m)$  is the matching cost function and  $F_{i,j}(max)$  is the maximum matching. We set the parameter  $T$  to 0.8 in this paper. This method considers location distribution and geometrical relationships among the feature points. Even in the case of few feature points, such as Fig 1A, it can achieve highly accurate matching, as shown in Fig 6.



**Fig 6. The matching result of Fig 1A by our proposed method.**

<https://doi.org/10.1371/journal.pone.0190383.g006>

To make the matching method more robust, we employ the progressive sample consensus algorithm that improves the RANSAC to remove outliers. The RANSAC algorithm does not deal well with the situation in which the number of mismatched pairs is too large in the proportion to the total matched pair, and it may fail, as shown in Fig 1C. The progressive sample consensus algorithm uses a data subset with a high matching rate as the sample set to estimate fitting [27]. It realizes rapid convergence and can deal with high mismatching. Following the methods discussed above, we can select feature points.

Then, the matched points can be easily sorted according to quality. The steps to remove the final mismatched pairs in this paper are as follows.

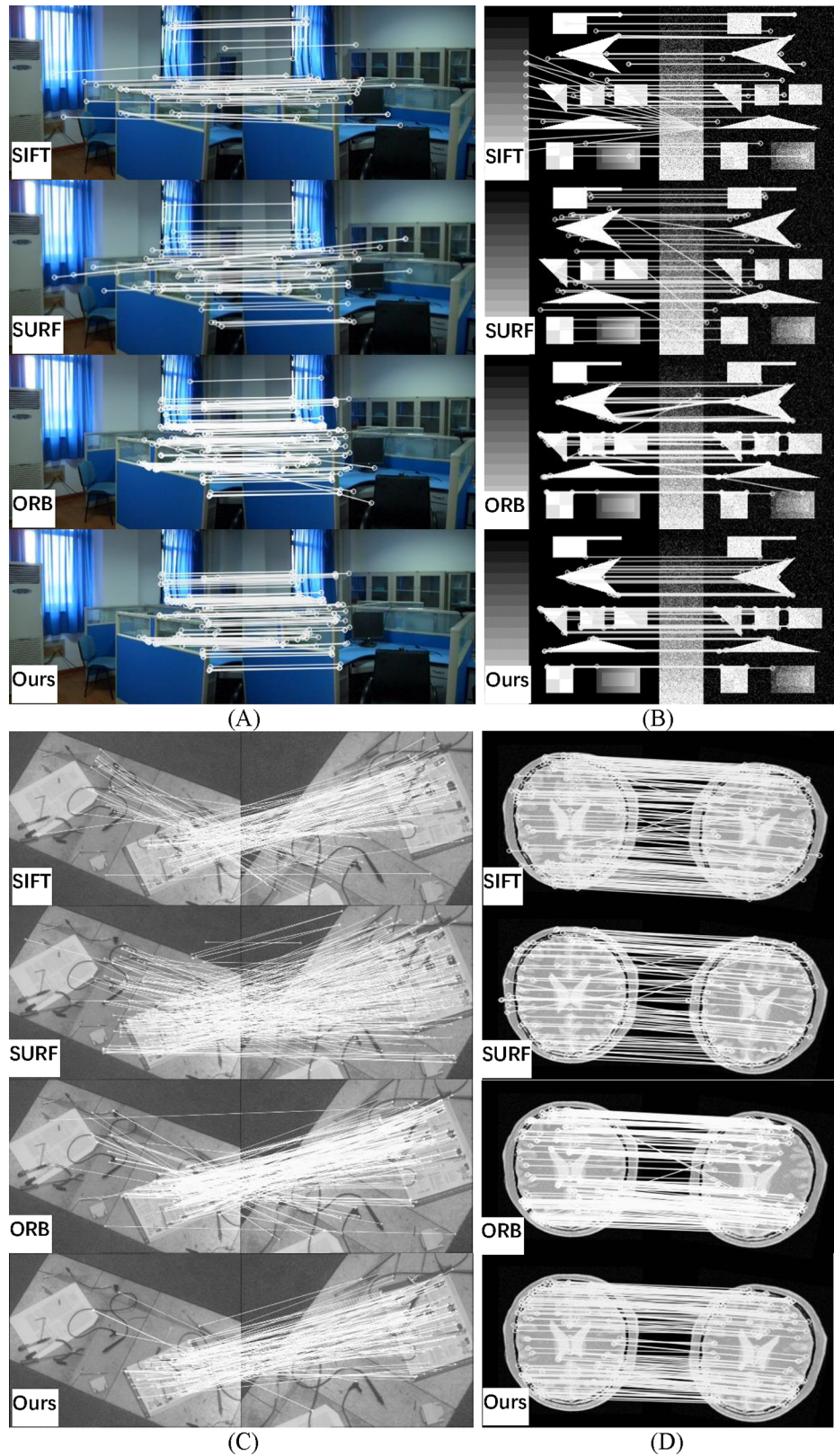
- 1) Feature points, identified following the procedures discussed above, are sorted in descending order by matched degree.
- 2) Set the sampling frequency, sample set  $\lambda$ , and sample size  $\eta$ . Each sample includes the coordinates of one feature point. The initial sample size  $\eta = 4$  is the minimum to estimate the transformation matrix. After each loop, the sample size increases to  $\eta = \eta + 1$ .
- 3) Determine the initial 4 sample set. Three matching point pairs are randomly chosen from  $\lambda$ , and then the  $\eta$ th matching point pair from  $\lambda$  is added to constitute the initial sample set. The transformation matrix  $\mathbf{M}$  transforms the coordinate from  $(u', v')$  into  $(u, v)$  and can be estimated from the initial sample set,

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{M} \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha & \Delta u \\ \sin\alpha & \cos\alpha & \Delta v \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix}. \quad (23)$$

- 4) Judge whether the initial matching point pair can satisfy the following two sampling termination conditions. If the conditions are not met, repeat step 2. Otherwise, exit the loop.
  - ① The ratio between the number of inner points and the total points is larger than the error threshold,  $\xi$ . If  $p$  and  $p'$  are a matching point pair, the condition to be an inner point is  $\|\mathbf{M}p - p'\|^2 \leq \xi$ .
  - ② The rate of increase of inner points should be less than the increase threshold,  $\xi'$ . In other words, the number of inner points should increase slowly after a certain number of samplings.

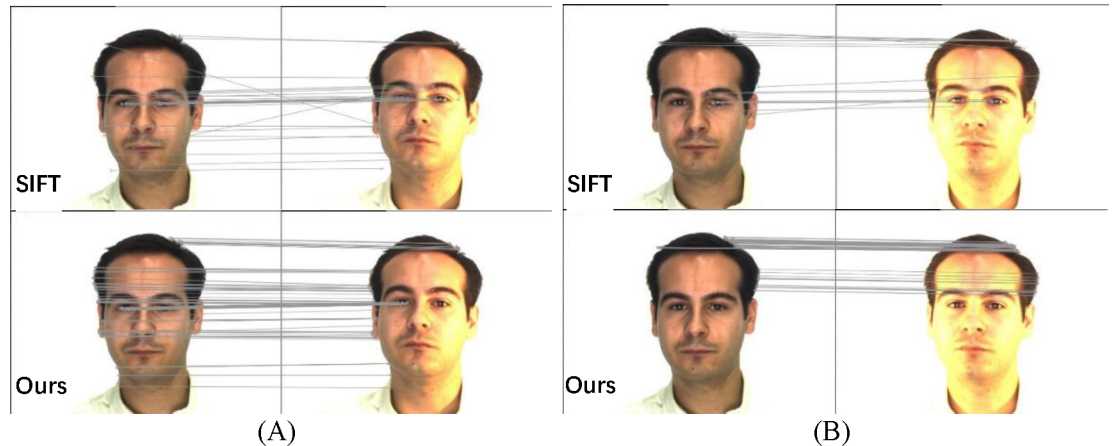
## Experimental procedure and results

The SIFT algorithm is often used in image registration and has better performance than many others methods. Here, we compare the SIFT algorithm to the BF algorithm employed to match the feature points and the RANSAC algorithm employed to eliminate mismatches. We show the final visual matching results. To examine the performance and robustness of our proposed method in various situations, test images were chosen with different resolutions and different applications. Fig 7, Fig 8, Fig 9, Fig 10 and Fig 11B are from the public image database. The other images were newly taken using a digital camera. Visual qualitative contrasts (including the connecting line between matching point pairs in two images) and quantitative comparisons (involving the different parameters and calculation results) were performed.



**Fig 7. Visual matching between SIFT, SURF, ORB and the proposed method for typical images. (A) indoor; (B) noise; (C) remote sensing; and (D) medical.**

<https://doi.org/10.1371/journal.pone.0190383.g007>



**Fig 8. Visual matching between SIFT and the proposed method for images with more difficult registrations. (A) small illumination variation; (B) strong illumination variation.**

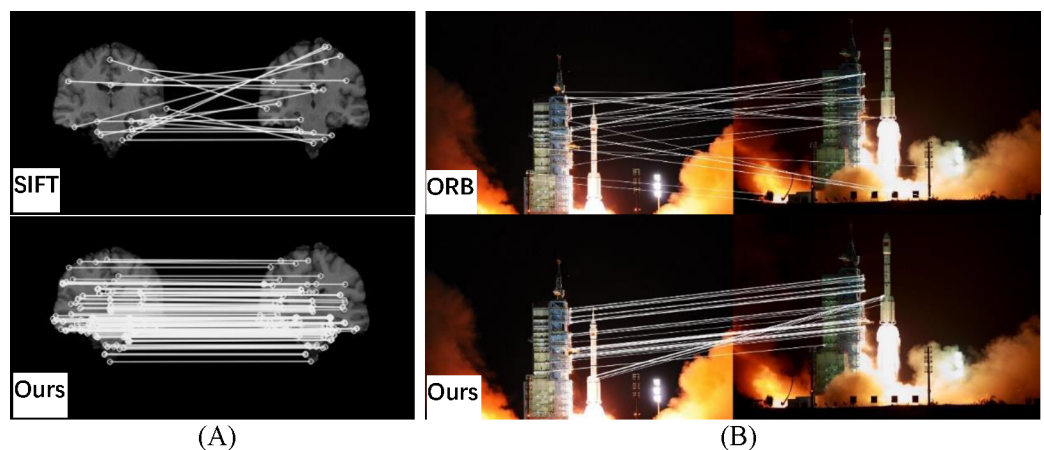
<https://doi.org/10.1371/journal.pone.0190383.g008>

### Qualitative comparison results

Fig 7 shows a visual comparison of typical images. Methods such as SIFT, SURF, ORB and the proposed method are compared in each group. The SIFT algorithm combined with RANSAC has good performance, but there are still some significant mistakes that are difficult to remove by using RANSAC. In our experiment of other registration algorithms, such as SURF and ORB, the results have similar outcomes to SIFT. All methods perform well when faced with a rotation of over 30 degrees, as shown in Fig 7C. The proposed method is significantly more accurate, and it provides a more uniform key point distribution due to the multi-layer selection.

Fig 8 shows the outcomes for more challenging images with excess exposure of the face. The images in Fig 8A and 8B are from the public Purdue AR face image database [36]. The SIFT algorithm produces many mismatches, whereas the proposed method has more reliable matching. Large numbers of correct matches are produced for varying illumination.

Fig 9 shows increasingly complex images. The images in Fig 9A have many similar details with a single background. This makes a large number of feature points difficult to match, and



**Fig 9. Visual matching among SIFT, ORB and the proposed method for more challenging images. (A) medical; (B) launch of the Tiangong rocket.**

<https://doi.org/10.1371/journal.pone.0190383.g009>

most of them are removed by RANSAC. The SIFT algorithm provides many mismatches, while the proposed method can precisely complete registration.

Fig 9B images are from the launch of the Tiangong rocket at different times. They have rich details, but the image content changes significantly. There are also large differences in brightness between the tail flame of the rocket and the dark background. The changing flight altitude and thick smoke can also seriously affect registration. Fig 1A and the upper image in Fig 9B are the result of SIFT and ORB, respectively. It shows that both the SIFT and ORB algorithms fail to accurately register this case. In contrast, the proposed method still has a large number of high quality matches.

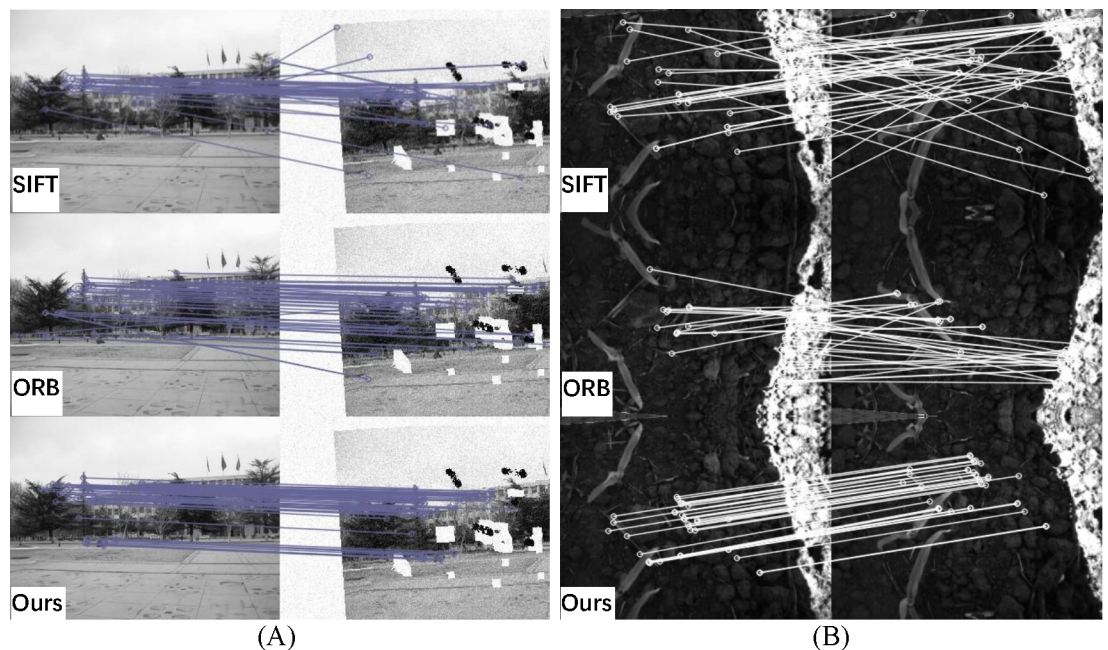
In summary, the proposed method shows excellent performance in feature extraction, selection and matching, even for very complex images.

## Quantitative comparison results

A further four groups of images were analyzed using SIFT and ORB with BF. Then, RANSAC was used to remove feature selection mismatches. These images correspond to different situations, including low signal-to-noise ratio, high dynamic range, lack of textures and medical images with local similarity. The number of matching points, the points after screening, the number of false matches and the matching rate were calculated and compared. Visible results and the quantitative comparisons for these test images are shown in Figs 10 and 11 and Table 1, respectively.

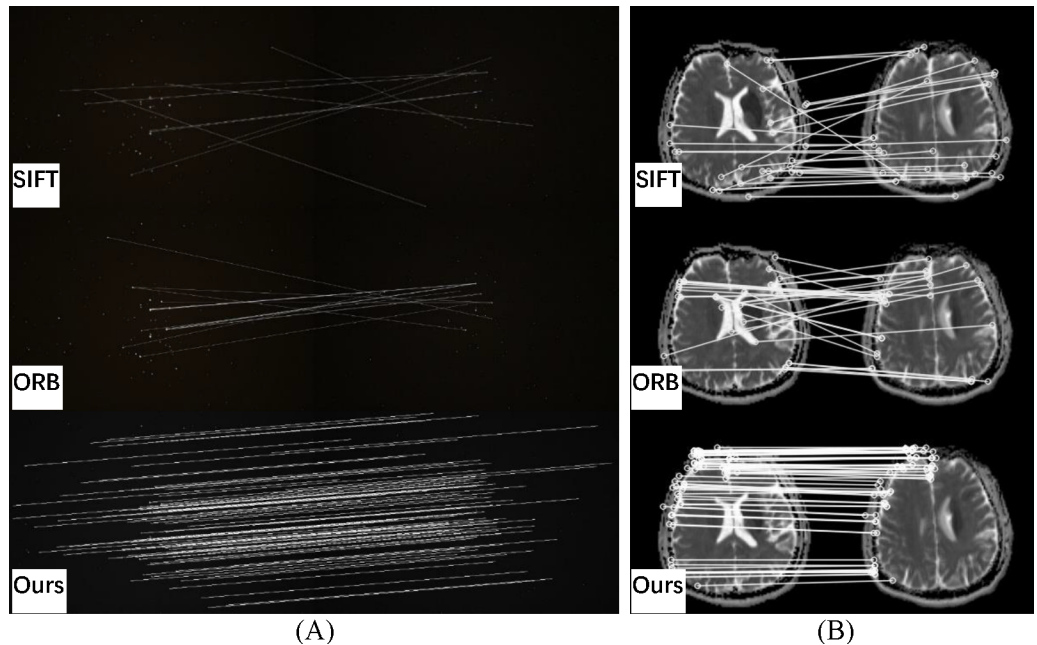
Fig 10A images show strong noise interference and low image quality. Both the SIFT and ORB algorithms have nearly half the error matches after screening by RANSAC. The proposed method removes the low reliability matching points, thus significantly reducing the final mismatch rate.

The images in Fig 10B have a large dynamic range and many similar features, but they lack details. Although a large number of feature points is extracted by SIFT and ORB, the methods



**Fig 10. Visual matching for SIFT, ORB, and the proposed method.** (A) low SNR (signal-to-noise ratio); (B) high dynamic range.

<https://doi.org/10.1371/journal.pone.0190383.g010>



**Fig 11. Visual matching for SIFT, ORB, and the proposed method.** (A) Pleiades; (B) medical.

<https://doi.org/10.1371/journal.pone.0190383.g011>

still suffer a high mismatch rate, and the algorithms are almost invalid. The proposed algorithm has outstanding performance in this situation, with a final matching rate of 97.06%, compared to 38.18% for the SIFT algorithm and 22.86% for the ORB algorithm.

Fig 11A shows the registration result of the Pleiades images. The exposure time of reference image (left) is 0.5 s, and the ISO is 6400. The exposure time of the target image (right) is 0.25 s,

**Table 1. Comparative metrics for the images in Figs 10 and 11.**

Image pair	Image type	Resolution	Methods	Initial matching points	Matching points after screening	Mismatches after screening	Matching rate
Fig 10A	low SNR ratio images	320×240	SIFT, RANSAC	468	38	16	57.89%
			ORB, RANSAC	511	51	23	54.90%
			<b>Proposed</b>	<b>556</b>	<b>73</b>	<b>7</b>	<b>90.41%</b>
Fig 10B	high dynamic range images	320×240	SIFT, RANSAC	569	55	34	38.18%
			ORB, RANSAC	391	35	27	22.86%
			<b>Proposed</b>	<b>771</b>	<b>68</b>	<b>2</b>	<b>97.06%</b>
Fig 11B	Pleiades images	749×249	SIFT, RANSAC	387	15	15	0.00%
			ORB, RANSAC	499	19	14	26.32%
			<b>Proposed</b>	<b>486</b>	<b>97</b>	<b>0</b>	<b>100.00%</b>
Fig 11B	medical images	230×230	SIFT, RANSAC	450	28	17	39.29%
			ORB, RANSAC	411	34	22	35.29%
			<b>Proposed</b>	<b>622</b>	<b>150</b>	<b>14</b>	<b>90.67%</b>

<https://doi.org/10.1371/journal.pone.0190383.t001>

**Table 2. Time consumption.**

Methods	Time to analyze the image group (s)				Average time (s)
	Fig 10A	Fig 10B	Fig 11A	Fig 11B	
SIFT	3.321	1.689	29.886	2.958	9.464
ORB	0.283	0.539	1.357	0.263	0.611
<b>Ours</b>	<b>0.346</b>	<b>0.591</b>	<b>1.557</b>	<b>0.322</b>	<b>0.704</b>

<https://doi.org/10.1371/journal.pone.0190383.t002>

and the ISO is 12800. The star images have almost no texture but a high dynamic range. The results show that the SIFT and ORB both fail to properly register this case, whereas the proposed method shows outstanding performance with a 100% matching rate. The images in Fig 11B were taken at different times with many similar details, and they have local deformation, which often occurs in medical images. The SIFT and ORB algorithms have a large number of false matches, so the RANSAC algorithm cannot work well, which can lead to a matching failure. The proposed method has a larger number of correct matches, with a matching rate of 90.67%.

The average time consumption of the different methods is shown in Table 2. Feature matching is the most time consuming aspect. If  $N_1$  and  $N_2$  are the number of feature points in two images, then the complexity of BF matching is  $O(N_1 * N_2)$ . Therefore, the number of key points involved in matching will directly impact the algorithm's efficiency. In general, in the step of feature detection, the number of key points is similar between SIFT/ORB and the proposed method, but the proposed method's processing can be up to 10 times faster than SIFT. In addition, the proposed method's processing speed is similar to that of ORB, but the number of correct matches is far greater than that of ORB.

## Conclusion and future work

Considering the limitations of traditional methods, this paper proposed a fast and robust image registration approach based on local features and geometric invariants. In the step of feature detection and description, we proposed an improved method of the ORB algorithm. The proposed method is scale invariant and produces more higher quality feature points. Then, we improved the removal of mismatches by combining it with the distribution of key points. A new distance constraint window is set according to the distribution of feature points, and the bidirectional matching constraint from the K nearest neighbor is utilized to extract higher quality feature points.

To further improve the method's adaptability and robustness and obtain the optimum matching point pairs, we proposed a novel geometric constraints matching algorithm with a new feature description vector based on the geometric invariance and a new cost function. Appropriate selection criteria were established to remove unreliable matches, and we integrated the PROSAC algorithm to further remove false matches. Thus, even with complex situations, we are still able to accurately register the images.

The experimental results show that our proposed registration method has superior adaptability and stronger robustness in terms of increasing the number of reliable key points and reducing the mismatch rate compared to the SIFT and ORB algorithms.

Future work may include improving the efficiency and real-time applications. The proposed method has fast matching speed using binary string descriptors, so parallel processing may be considered to allow higher resolution images to be processed in real time.

## Author Contributions

**Data curation:** Tinghua Zhang.



**Investigation:** Tinghua Zhang.

**Methodology:** Yan Lu.

**Project administration:** Kun Gao.

**Supervision:** Tingfa Xu.

**Validation:** Tingfa Xu.

**Visualization:** Yan Lu.

**Writing – original draft:** Yan Lu.

**Writing – review & editing:** Kun Gao.

## References

1. Wang X, Shen S, Ning C, Huang F, Gao H. Multi-class remote sensing object recognition based on discriminative sparse representation. *Appl Opt.* 2016; 55: 1381–1394. <https://doi.org/10.1364/AO.55.001381> PMID: 26906591
2. Li S, Li H, Zheng Z, Peng Y, Wang S, Liu X. Full-parallax three-dimensional display using new directional diffuser. *Chin Opt Lett.* 2011; 9: 081202.
3. Mendelowitz S, Klapp I, Mendlovic D. Design of an image restoration algorithm for the TOMBO imaging system. *J Opt Soc Am A Opt Image Sci Vis.* 2013; 30: 1193–1204. <https://doi.org/10.1364/JOSAA.30.001193> PMID: 24323107
4. Can T, Karali AO, Aytac T. Detection and tracking of sea-surface targets in infrared and visual band videos using the bag-of-features technique with scale-invariant feature transform. *Appl Opt.* 2011; 50: 6302–6312. <https://doi.org/10.1364/AO.50.006302> PMID: 22108891
5. Chen J, Luo L, Liu C, Yu JG, Ma J. Nonrigid registration of remote sensing images via sparse and dense feature matching. *J Opt Soc Am A Opt Image Sci Vis.* 2016; 33: 1313–1322. <https://doi.org/10.1364/JOSAA.33.001313> PMID: 27409688
6. Wang R, Xia Y, Wang G, Tian J. License plate localization in complex scenes based on oriented FAST and rotated BRIEF feature. *J Electron Imaging.* 2015; 24: 053011.
7. Hou W, Zhu J, Yang T, Jin G. Construction method through forward and reverse ray tracing for a design of ultra-wide linear field-of-view off-axis freeform imaging systems. *J Opt.* 2015; 17: 055603.
8. Zhu J, Wang L, Yang R, Davis JE, Pan Z. Reliability fusion of time-of-flight depth and stereo geometry for high quality depth maps. *IEEE Trans Pattern Anal Mach Intell.* 2011; 33: 1400–1414. <https://doi.org/10.1109/TPAMI.2010.172> PMID: 20820074
9. Oliveira FP, Tavares JM. Medical image registration: a review. *Comput Methods Biomech Biomed Engin.* 2014; 17: 73–93. <https://doi.org/10.1080/10255842.2012.670855> PMID: 22435355
10. Zhang J, Yue M, Liu H. Dynamic PET image reconstruction with Geometrical structure prior constraints. *J Zhejiang Uni (Eng Sci).* 2012; 46: 961–966.
11. Feng B, Chen F, Liu G, Xiang Y, Liu B, Lv Z. Image-based displacement and rotation detection using scale invariant features for 6 degree of freedom ICF target positioning. *Appl Opt.* 2015; 54: 4130–4134.
12. Moravec HP. Towards automatic visual obstacle avoidance. *Proceedings of international joint conference on artificial intelligence.* Cambridge, MA, USA; 1997. pp. 584–590.
13. Li Y, Wang S, Tian Q, Ding X. A survey of recent advances in visual feature detection. *Neurocomputing.* 2015; 149: 736–751.
14. Lowe DG. Distinctive image features from scale-invariant keypoints. *Int J Comput Vis.* 2004; 60: 91–110.
15. Chen M, Shao Z, Li D, Liu J. Invariant matching method for different viewpoint angle images. *Appl Opt.* 2013; 52: 96–104. PMID: 23292380
16. Bay H, Tuytelaars T, Van Gool L. SURF: speeded up robust features. In: Leonardis A, Bischof H, Pinz A, editors. *Computer vision—ECCV 2006.* Berlin, Heidelberg: Springer; 2006. pp. 404–417.
17. Rosten E, Drummond T. Machine learning for high-speed corner detection. In: Leonardis A, Bischof H, Pinz A, editors. *Computer vision—ECCV 2006.* Berlin, Heidelberg: Springer; 2006. pp. 430–443.
18. Tayara H, Ham W, Chong KT. A real-time marker-based visual sensor based on a FPGA and a soft core processor. *Sensors.* 2016; 16: 2139.

19. Calonder M, Lepetit V, Strecha C, Fua P. BRIEF: binary robust independent elementary features. In: Daniilidis K, Maragos P, Paragios N, editors. *Computer vision—ECCV 2010*. Berlin, Heidelberg: Springer; 2010. pp. 778–792.
20. Rublee E, Rabaud V, Konolige K, Bradski G. ORB: an efficient alternative to SIFT or SURF. *IEEE International conference on computer vision*. Barcelona, Spain; 2011. pp. 2564–2571.
21. Zhang J, Chen L, Wang X, Teng Z, Brown AJ, Gillard JH, et al. Compounding local invariant features and global deformable geometry for medical image registration. *PLoS One*. 2014; 9: e105815. <https://doi.org/10.1371/journal.pone.0105815> PMID: 25165985
22. Kahaki SMM, Nordin MJ, Ashtari AH, Zahra SJ. Invariant feature matching for image registration application based on new dissimilarity of spatial features. *PLoS One*. 2016; 11: e0149710. <https://doi.org/10.1371/journal.pone.0149710> PMID: 26985996
23. Li Y, Jin H, Qiao W, Jing J, Yu H. Robustly building keypoint mappings with global information on multi-spectral images. *EURASIP J Adv Signal Process*. 2015; 2015: 53.
24. Lee DH, Lee DW, Han BS. Possibility study of scale invariant feature transform (SIFT) algorithm application to spine magnetic resonance imaging. *PLoS One*. 2016; 11: e0153043. <https://doi.org/10.1371/journal.pone.0153043> PMID: 27064404
25. Jia K, Chan T-H, Zeng Z, Gao S, Wang G, Zhang T, et al. ROML: a robust feature correspondence approach for matching objects in a set of images. *Int J Comput Vis*. 2016; 117: 173–197.
26. Bastanlar Y, Temizel A, Yardimci Y. Improved SIFT matching for image pairs with scale difference. *Electron Lett*. 2010; 46: 346–348.
27. Fischler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun ACM*. 1981; 24: 381–395.
28. Chum O, Matas J. Matching with PROSAC- progressive sample consensus. *IEEE computer society conference on computer vision and pattern recognition*. Washington, DC; 2005. pp. 220–226.
29. Kapela R, Gugala K, Sniatala P, Swietlicka A, Kolanowski K. Embedded platform for local image descriptor based object detection. *Appl Math Comput*. 2015; 267: 419–426.
30. Yang S, Zhang J, Zhang W. Phase-sensitive periodical correlation of local beam descriptors for image registration. *Neurocomputing*. 2016; 173: 1694–1705.
31. Liu C, Ma J, Ma Y, Huang J. Retinal image registration via feature-guided Gaussian mixture model. *JOSA A*. 2016; 33: 1267–1276. <https://doi.org/10.1364/JOSAA.33.001267> PMID: 27409682
32. Wu X, Zhao Q, Bu W. A SIFT-based contactless palmprint verification approach using iterative RAN-SAC and local palmprint descriptors. *Pattern Recognit*. 2014; 47: 3314–3326.
33. Wang B, Lu Q, Li Y, Li F, Bai L, Lu G, et al. Image registration method for multimodal images. *Appl Opt*. 2011; 50: 1861–1867. <https://doi.org/10.1364/AO.50.001861> PMID: 21532665
34. Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. *IEEE Trans Pattern Anal Mach Intell*. 2002; 24: 509–522.
35. Liao D, Wang P, Zhao J, Gregersen H. Validation of shape context based image registration method using digital image correlation measurement on a rat stomach. *J Comput Med*. 2014; 2014: 504656.
36. Martinez AM, Benavente R. The AR face database. *CVC Technical Report*. 1998. pp. 24.