IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 5 April 2023, accepted 23 April 2023, date of publication 27 April 2023, date of current version 3 May 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3270911*

RESEARCH ARTICLE

# An SIFT-Based Fast Image Alignment Algorithm for High-Resolution Image

**ZETIAN TANG[1], ZEMIN ZHANG[2], WEI CHEN[3], AND WENTAO YANG[1]**
[1]School of Physics and Electrical Engineering, Liupanshui Normal University, Liupanshui 553004, China
[2]School of Chemistry and Materials Engineering, Liupanshui Normal University, Liupanshui 553004, China
[3]School of Electronics and Communication Engineering, Sun Yat-sen University, Shenzhen 518107, China

Corresponding author: Zetian Tang (tang_zetian@foxmail.com)

This work was supported in part by the Natural Science Research Project of Guizhou Provincial Education Office under Grant Qianjiaohe KY Zi [2022] 046 and Grant Qianjiaohe KY Zi [2020] 129.

**ABSTRACT** To solve the problem of high-resolution image alignment time overhead, an SIFT-based fast image alignment algorithm is presented. The overlap region of images is computed in detail by phase correlation algorithm to avoid a lot of useless calculations of non-overlapping region. After the distribution of feature points determined in difference of Gaussian through formula derivation, the total number of feature points is limited. The more stable spatially distributed for the feature points is obtained due to the expanded detection range of extreme points and added non-maximum suppression. It is noteworthy that the range of the descriptor is calculated by the method of down-sampling. And the circular descriptor is constructed with only 56-dimensional in the feature point descriptor generation stage, which makes the time of the descriptor generation and feature point matching shorter. This indicates that the total descriptor calculation is faster in lower dimensions by the new algorithm. In addition, experimental results show that the average time (9.60s, 13.46s, and 15.81s) of the proposed algorithm is 0.86%, 0.43%, and 0.10% of the SIFT algorithm, respectively. The overall speed is 2-3 orders of magnitude faster than the SIFT algorithm, which indicates that the new algorithm can solve the problem of high-resolution image alignment time overhead. The new algorithm provides a good stitching quality and shows an excellent performance for high-resolution image compared with several existing image stitching algorithms at the current. It indicates that the algorithm has potential application value in real-time image stitching.

**INDEX TERMS** Scale-invariant feature transform (SIFT), image alignment, high-resolution image, fast alignment.

## I. INTRODUCTION

In recent years, image stitching technology is widely used in remote sensing, motion detection, resolution enhancement, medical imaging and other fields [1], [2], [3], [4], [5]. As the important part of the image processing field, image stitching algorithms are consisted of two main steps: image alignment and image blending. The development of image stitching technology typically depends on the innovations of these two aspects. Image alignment is used to obtain the motion relationship by detecting and matching feature points across two images or multiple images. It directly associates with the speed and success rate of image stitching process [6].

At present, different feature detection-description algorithms are emerged, such as Harris, oriented FAST and rotated BRIEF (ORB) [7], binary robust invariant scalable keypoints (BRISK) [8], scale-invariant feature transform (SIFT) [9] and speed up robust features (SURF) [10]. The Harris and ORB algorithms have good speed, but no scale invariance. They can obtain scale invariance through Gaussian pyramid [11], [12]. Although BRISK algorithm has well rotation and scale invariance, the overall time is longer than ORB algorithm. SIFT algorithm has well effects in translation, rotation, illumination, scaling and affine transformation, but its computational complexity is large. SURF-based

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

mosaicism techniques are faster than SIFT-based techniques. However, they perform poorly under certain variations (particularly color, illumination, some affine transformation) [1]. Some research combines the SURF algorithm with other algorithms to improve its performance [13], [14].

Among many of image stitching algorithms, SIFT algorithm has won the favor of many researchers because of its excellent performance and good robustness. In terms of quality improvement of image stitching, Laraqi [15] and Yan et al. [16] have improved the image stitching quality through image preprocessing in recent years. Gong [17] and Guang et al. [18] have achieved the matching from rough to fine used two-step matching strategy. Chang [3] and Zhao et al. [19] have improved the quality of image stitching by applied of a new matching method. Ma et al. [2] have enhanced the matching of feature points through combined with the gradient definition method and information of key points. Gong et al. [20] proposed an invariant descriptor of robust neighborhood structure and designed a strategy of dynamic matching.

The resolution of image is gradually improved with the improvement of image acquisition equipment performance. High-resolution images have more information, and it is important that how to use the information from high-resolution images. Using high resolution image stitching can preserve more details of the image, while it also means greater computational complexity. It is becoming increasingly prominent that how to solve the problem of large computational complexity of high-resolution images. However, in terms of image stitching efficiency, most existing algorithms focus on the improvement of low-resolution image stitching and image stitching at a certain stage. Zhang et al. [21] and Zhao et al. [22] have changed the contrast threshold of feature point extraction. Zhang [11], Zhao [22], Kupfer [23] and Ma et al. [24] have improved the matching method of image stitching and the efficiency of algorithm. Zeng et al. [25] have realized real-time adaptive registration of visible and infrared images through used morphological gradient and C_SIFT. Chen et al. [26] have reduced heavily the time cost in the matching phase of SIFT algorithm combined with the SIFT algorithm Canny edge detection and the new descriptor whose size (only 18-dimensional). Shi et al. [27] have achieved a faster speed and better splicing quality by following improvements: They divided the image into feature blocks by the improved fuzzy C-Means (FCM) algorithm, extracted the image feature descriptors through the SIFT algorithm and avoided ghosting and shape distortion by optimized the overlapping regions. The phase correlation algorithm and texture complexity classification of the image have been used by Wang et al. [28] to detect feature points only in regions of high complexity, which reduced the regions detected. Xu et al. [12] improved the performance and precision of corners extracted by Harris, and then used the SIFT operator to speed up the registration process. Zhou et al. [29] have reduced the dimension of SIFT descriptors to reduce the dimension of feature descriptor by used convolutional neural networks. Through Li and coworker's [30] efforts, the algorithm running time of useless regions have been reduced by divided regions of the similarity of shared information between two images. They have also improved the quality of image stitching through used two projections stitching during the stitching phase to reduce algorithm ghosting that can make the algorithm extract feature points more efficiently. Du et al. [31] have achieved good accuracy and efficiency due to the spatial transformation images modeling and introduced of the robust Bayesian framework. In addition, the global information and descriptors were used to establish key point mapping by Li et al. [32]. They have also limited the matching space of feature points at the low-level pyramids started matching from the high-level pyramids to achieve the matching from rough to fine. These researches have reduced effectively the matching time of the SIFT algorithm.

However, although image stitching has been made good progress, there are few researches in the field of high-resolution images. It is becoming increasingly prominent that how to solve the problem of large computational complexity of high-resolution images. We built three datasets of high-resolution images and tested them with several existing image stitching algorithms at the current. The experimental results show that the time cost is too large to meet the real-time requirements. It can be seen that although the existing algorithms have achieved good performance in the field of fast stitching of low-resolution images, their performance is poor when using high-resolution image stitching. To solve this problem, a new fast alignment algorithm for high-resolution image is presented based on SIFT through following three aspects:

(1) The phase correlation algorithm is applied to calculate the similar regions of image to avoid a lot of useless calculations of non-overlapping regions. The distribution of feature points is determined in the Difference of Gaussian (DOG) through formula derivation. And the influence of different number of feature points is analyzed for image stitching quality. After that, the number of feature points is selected for that is conducive to image stitching speed and quality.

(2) To improve the stability of feature points, the range of extreme point detection is expanded and the influence of different detection ranges is analyzed on feature points. In order to improve the spatial distribution of feature points and the utilization of image information, non-maximum suppression (NMS) is added in the process of extreme point detection, and the influence of different NMS ranges is analyzed on feature points. It is selected that for the suitable detection range of extreme points and NMS range.

(3) A circular feature descriptor is designed with 56-dimensional to solve the problem that the poor robustness and large dimension of SIFT descriptor. Besides, we consider a larger calculation range for the descriptor, and the calculation regions of the descriptor are down-sampled to give consideration to both speed and accuracy. Results show that

our proposed algorithm provides a better stitching quality and shows an excellent performance for high-resolution image compared with several existing image stitching algorithms at the current, which indicates that the algorithm has potential application value in real-time image stitching.

## II. RELATED WORK

As the most popular algorithm in image stitching field, SIFT algorithm has good performance and robustness. It can be divided into the following four steps: (1) Detection of spatial extreme points; (2) Removing unstable edge response points and extreme points of poorer contrast; (3) Calculating the direction of key points; (4) Descriptor generation.

### A. DETECTION OF SPATIAL EXTREME POINTS

Initially, a scale space is constructed by convolving an image repeatedly using a Gaussian filter with changing scales and grouping the outputs into octaves as:

$$L(x, y, \sigma) = G(x, y, \sigma) \otimes I(x, y) \tag{1}$$

where, $I(x, y)$ is the input image, $G(x, y, \sigma)$ is the Gaussian filter. The Gaussian filter is defined as follows:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-m/2)^2+(y-n/2)^2}{2\sigma^2}} \tag{2}$$

where, $\sigma$ is the standard deviation of the normal distribution, $m$ and $n$ are the size of the Gaussian filter, and $x$ and $y$ are the positions of the corresponding elements of the Gaussian filter.

The spatial extreme points are detected by the DOG, which is obtained by subtracting between two adjacent layers of an octave of Gaussian pyramids:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{3}$$

where, $L(x, y, k\sigma)$ from formula (1), $k$ represents the scale factor of two scale spaces. The DOG model is shown in figure 1 [9].

The local extreme points are searched as candidates for feature points after of establishment DOG. The detection method of local extreme points is mainly comparing any point $(x, y)$ with 8 adjacent points on the same layer, the 9 points on the upper layer and lower layer respectively. The point is a candidate for the feature point when the value of point $(x, y)$ is the maximum or minimum. The key points of DOG can be extracted in the middle layer (except the top layer and the bottom layer) of each octave.

### B. REMOVING UNSTABLE EDGE RESPONSE POINTS AND EXTREME POINTS OF POORER CONTRAST

The extreme points of low and the edge response points of poor stability contrast need to be further removed after the extreme points are obtained, which can improve the anti-noise ability and stability of the extreme points.

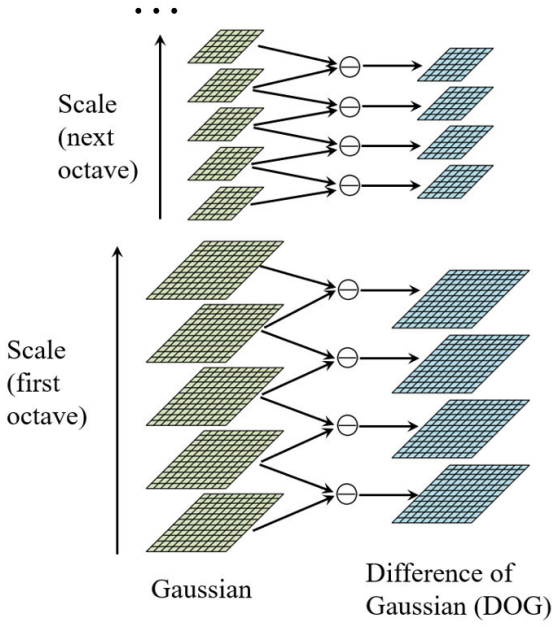


**FIGURE 1.** DOG model.

Calculate the $D(\hat{X})$ value of the extreme point, and the calculation formula is as follows:

$$D(\hat{X}) = D + \frac{1}{2}\frac{\partial D^T}{\partial X_0}\hat{X} \tag{4}$$

where, $\hat{X} = (x, y, \sigma)^T$ denotes the center offset of the relative interpolation, $D$denotes the first term of the Taylor expansion of the spatial scale function $D(x, y, \sigma)$ at the extreme point. Then, all extreme with a value of $D(\hat{X})$ less than 0.03 were discarded to obtain more stable extreme values.

The location and scale of a feature point can be pinpointed accurately by a Hessian matrix, which can be expressed as:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \tag{5}$$

Then, the stability of the point is presented by the following formula:

$$stability = \frac{(D_{xx} + D_{yy})^2}{D_{xx}D_{yy} - D_{xy}^2} < \frac{(r+1)^2}{r} \tag{6}$$

where, $r$ represents the parameter which controls feature value. Remove all points that do not conform to Formula (6).

### C. CALCULATING THE DIRECTION OF KEY POINTS

For the SIFT operator to be rotated invariant, the main direction of the key points needs to be determined by the following formula: (7) and (8), as shown at the bottom of the next page, where: $L$ represents the scale space value of the key point, and formulas (7) and (8) are the corresponding gradient modulus value and direction, respectively. The direction of the maximum is the main direction. In addition, they are set as secondary directions to the key points if the directions

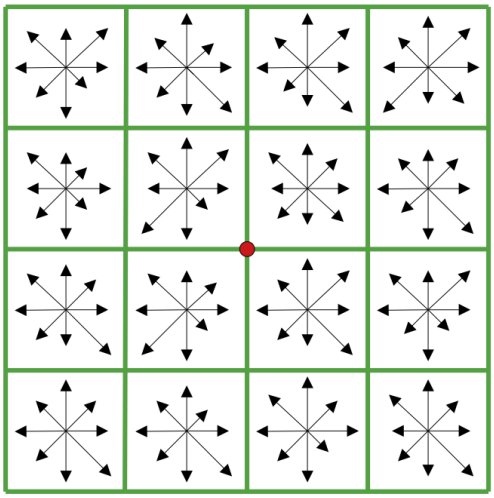

**FIGURE 2.** Descriptor of SIFT.

are greater than 80% of the maximum. Each direction is preserved as a feature point, and the descriptors of the feature points are calculated separately.

### D. DESCRIPTOR GENERATION

The pixels around the feature points are rotated to the corresponding direction and the region is divided into $4 \times 4$ to ensure the invariance of the descriptor direction. As shown in figure 2, the values of cumulative gradient are calculated in eight directions of each region to form a 128-dimensional feature descriptor.

## III. PROPOSED METHOD

The works of this paper are mainly improving the SIFT algorithm to reduce the huge time cost of high-resolution image stitching. The overall block diagram is shown in figure 3. The overlapping regions of the reference image and the registration image are roughly determined using the phase correlation algorithm to reduce the computation of the non-overlapping regions. Then, in the feature point detection stage, the number of extracted feature points is reduced on the premise that ensuring a good effect of matching. In addition, expanded detection range of extreme points and added NMS to improve the stability of feature points and the utilization rate of image information. Subsequently, a 56-dimensional circular descriptor is established in the descriptor generation stage. The calculation range of the descriptor is down-sampled to further improve the calculation efficiency. Finally, the projection transformation matrix is calculated through feature point matching to complete image stitching.

### A. PHASE CORRELATION ALGORITHM

It is essential that to calculate the whole image for SIFT algorithm. The SIFT algorithm will produce some invalid calculations due to the provider of incompletely effective information regions. Therefore, the phase correlation algorithm is applied to preliminarily determine the overlapping region before the SIFT algorithm processing.

The algorithm first transforms the image to the frequency domain through Fourier transform, and then the shift parameters of the two images are calculated using the normalized cross power spectrum.

If there is shift $(x_0, y_0)$ in part of the information in the two images, that is,

$$f_2(x, y) = f_1(x - x_0, y - y_0) \tag{9}$$

Then, $f_1(x, y)$ and $f_2(x, y)$ are transformed to the frequency domain using Fourier transform to get $F_1(\mu, v)$ and $F_2(\mu, v)$, in which case the relationship between $F_1$ and $F_2$ is as follows:

$$F_2(\mu, v) = e^{-j(\mu x_0 - v y_0)} F_1(\mu, v) \tag{10}$$

The cross-power spectra of $F_1$ and $F_2$ is:

$$\frac{F_1^*(\mu, v) F_2(\mu, v)}{\left|F_1^*(\mu, v) F_2(\mu, v)\right|} = e^{-j(\mu x_0 - v y_0)} \tag{11}$$

Inverse Fourier transform is performed on the right side of formula (11) to obtain the impact response function $\delta(x\text{-}x_0, y\text{-}y_0)$. The range of the overlapping region can be preliminarily obtained by searching the maximum point $(x_0, y_0)$ of $\delta$. There $(x_0, y_0)$ is the optimal translation amount of the overlapping region between the images $f_1(x, y)$ and $f_2(x, y)$.

### B. FEATURE POINT DETECTION

The region detected by the traditional SIFT algorithm is only $3 \times 3 \times 3$. However, the image size of the high-resolution image is very large. The stability of feature points is poor due to the small detection range of feature points. The huge number of feature points extracted could lead to excessive calculation time in feature point detection, descriptor generation and matching stage, respectively. In addition, too many feature points would increase little the quality of image stitching. So, it is particularly important to select appropriate and effective feature points, which should be reasonably distributed in the DOG.

$$m(x, y) = = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \tag{7}$$

$$\theta(x, y) = \arctan\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \tag{8}$$

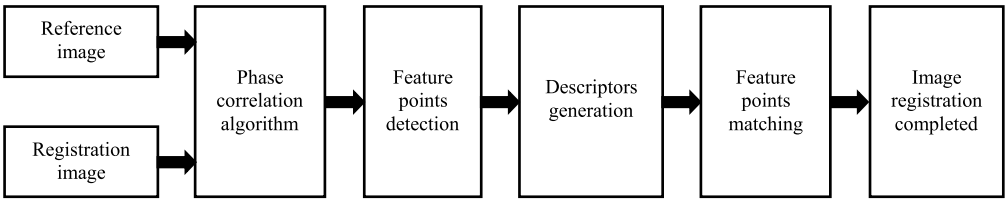


**FIGURE 3.** Block diagram of proposed method.

Two improvements are made in the phase of feature point detection: (1) The number of feature points is limited to improve the efficiency of the algorithm on the premise of guaranteeing a good matching result. (2) The extreme points are detected from a larger range ($7 \times 7 \times 3$) to ensure better stability of the extracted feature points. The extraction range of feature points is discussed in the next section. Additionally, NMS is used to prevent feature points from clustering in complex texture regions. The application of NMS can also improve the spatial distribution of feature points and the utilization of image information.

### 1) NUMBER OF FEATURE POINTS EXTRACTED

It is necessary to determine the number of feature points of each layer in each octave when limiting the total number of feature points. We have carried out mathematical derivation for this problem, and the process is as follows:

Firstly, set the total number of feature points to be $N_{sum}$, which is equal to the sum of the number of feature points of each layer of each octave:

$$N_{sum} = \sum_{o=1}^{m} \sum_{l=2}^{n-1} N_{o,l} \tag{12}$$

where, $m$ and $n$ are the number of octave and the number of layers in DOG, and $N_{o,l}$ is the number of feature points in the layer $l$ of octave $o$.

We set the ratio of the number of feature points extracted from the same layer of the two adjacent octaves to be $\alpha$. And set the extraction ratio of feature points from adjacent two layers to be $\beta$. Additionally, the number of feature points of the second layer is also set in the first octave as $N_{1,2}$, and the number of feature points for any layer in the DOG can be obtained:

$$N_{o,l} = N_{1,2}\alpha^{o-1}\beta^{l-2} l \geq 2 \tag{13}$$

The number of feature points is distributed proportionally in different layers of each octave and in the same layer of different octaves. The sum of the number of feature points for any layer in different octave as:

$$N_{sum,l} = \frac{N_{1,2}\beta^{l-2}(1-\alpha^o)}{1-\alpha} \; l \geq 2, \alpha \neq 1 \tag{14}$$

Because the Gaussian pyramid is obtained by the method of down-sampling, the proportion of feature point extraction cannot be 1 thus it is not considered $\alpha = 1$. $N_{sum,l} = \{N_{sum,2}, N_{sum,3}, \ldots N_{sum,n-1}\}$ is proportional distribution. Thus, the total number of feature points is:

$$N_{sum} = \begin{cases} (n-2)\sum_{o=1}^{m} \frac{N_{1,2}\beta^{-1}(1-\alpha^o)}{1-\alpha} & \alpha \neq 1, \beta = 1 \\ N_{1,2}\frac{(1-\alpha^o)(1-\beta^l)}{(1-\alpha)\beta(1-\beta)} & l \geq 2, \alpha \neq 1, \beta \neq 1 \end{cases} \tag{15}$$

Therefore, $N_{1,2}$ is

$$N_{1,2} = \begin{cases} \frac{N_{sum}}{(n-2)\sum_{o=1}^{m}\frac{1-\alpha^o}{1-\alpha}} & \alpha \neq 1, \beta = 1 \\ N_{sum}\frac{\beta(1-\alpha)(1-\beta)}{(1-\alpha^o)(1-\beta^l)} & l \geq 2, \alpha \neq 1, \beta \neq 1 \end{cases} \tag{16}$$

Take Formula (16) into Formula (13) to get:

$$N_{o,l} = \begin{cases} \frac{N_{sum}\alpha^{o-1}}{(n-2)\sum_{o=1}^{m}\frac{1-\alpha^o}{1-\alpha}} & \alpha \neq 1, \beta = 1 \\ N_{sum}\frac{(1-\alpha)(1-\beta)\alpha^{o-1}\beta^{l-1}}{(1-\alpha^o)(1-\beta^l)} & l \geq 2, \alpha \neq 1, \beta \neq 1 \end{cases} \tag{17}$$

The distribution of feature points can be calculated according to formula (17) after determining the total number.

The feature points are extracted from large to small pairs according to the contrast value after the number of feature points of each layer is obtained. It is necessary to extract them one by one and stop extraction when the limited number is reached due to the multiple secondary directions of each feature point and each secondary direction serves as a separate feature point. The number of extracted feature points may be slightly more than the limited number due to the secondary direction of feature points. The total number of feature points and the proportion of feature point distribution ($\alpha$ and $\beta$) are discussed in the feature point extraction of the next section.

### 2) FEATURE POINT EXTRACTION METHOD

Detection range of the descriptor is changed to $7 \times 7 \times 3$ to improve the stability of feature points. NMS is added to the detection of spatial extreme points to prevent feature points from clustering in complex textured regions and make the distribution of feature points more reasonable. It is not feasible to use a fixed pixel value suppression method since the size of each image is different. Therefore, the NMS range is set according to the size of the image after phase correlation calculation, and the range is as follows:

$$\{(x, y)|x \in (x_i - wM_p, x_i + wM_p), \\ y \in (y_i - wN_p, y_i + wN_p)\} \tag{18}$$

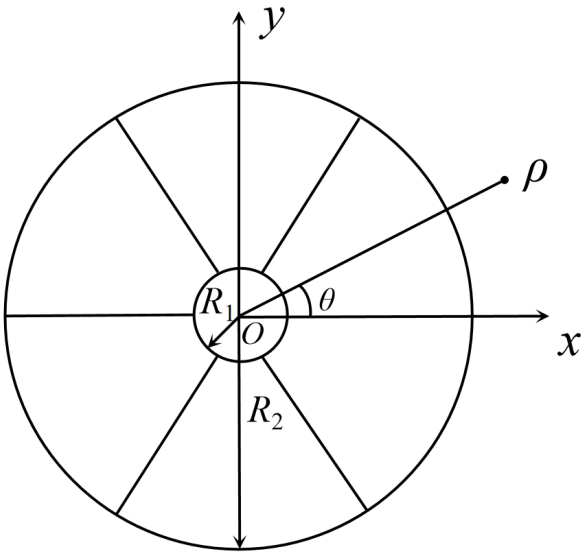


**FIGURE 4.** Descriptor structure.

where, $x_i$ and $y_i$ are the coordinates of the feature points, $M_p$ and $N_p$ are the size of the image detected by the phase correlation algorithm, $w$ is the proportion of NMS rang.

The extreme points are detected according to the contrast size first in spatial extreme point detection. NMS is performed in the region shown in Formula (18) after each feature point is extracted. Since there are multiple octave and layers of DOG, each octave uses an independent NMS. The NMS range is discussed in the next section.

### C. FEATURE DESCRIPTOR

The 128-dimensional descriptor of the SIFT algorithm leaded to the much time cost of feature point matching since. To address this issue, we use a GLOH-like [33.34] circular neighborhood (radius of $10\sigma$ and $12\sigma$) and log-polar sectors (7 location bins) to create a feature descriptor. A series of experiments shows that GLOH obtains the best results [33]. The descriptor is shown in figure 4 and the steps to create a descriptor are as follows: Establish double concentric circles with feature point as pole point and $R_1$, $R_2$ as polar radius, as well as outer concentric circles are divided into six equal parts. The ratio of $R_1$ to $R_2$ is 0.3. Then, the gradient cumulative values for eight directions are calculated in each region (the calculated results are used as descriptors). The discussion of the setting of $R_1$ and $R_2$ ratio and the division of peripheral ring is described in detail in the feature point descriptor of the next section.

The calculation range of the descriptor is small due to the large size of the high-resolution image. It is difficult for the descriptor to effectively describe the information near the feature points, which could result in inaccurate feature point matching. Therefore, it is necessary to increase the calculation range of descriptors. However, the increasing the calculation range of descriptors would lead to the complexity of calculation and time cost. To solve the above problem, expanded

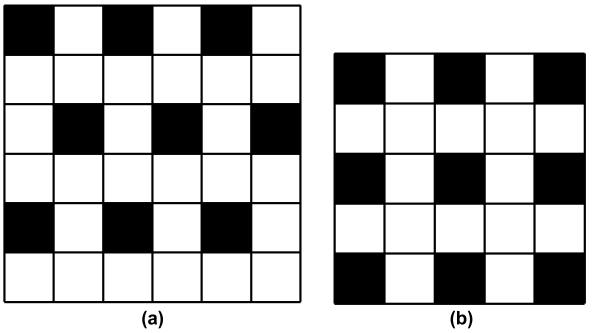


**FIGURE 5.** Descriptor down-sampling method: (a) The calculation region is an even number column; (b) The calculation region is an odd number column.

the calculation region and the method of down-sampling is applied for the calculated region before establishing the descriptor. The method of down-sampling (figure 5 (a)) is adopted when the columns of the feature point calculation region are even. The method of down-sampling (figure 5 (b)) is adopted as well when the number of columns is odd in the feature point calculation region. The down-sampled image is obtained through retained the pixel value of the black region and discarded the pixel value of the white region. The impact of the calculation range and down-sampling method are analyzed in the feature point descriptor of the next section.

## IV. EXPERIMENT RESULTS AND DISCUSSION

### A. DATASETS

We collected some high-resolution images and three datasets are established to evaluate the proposed algorithm.

(1) Dataset 1: The dataset contains 150 images of mobile phones and digital cameras. The images include scenes of buildings, mountains, cities, rivers, farmland, and so on. These images have rigid or affine transformations. The size of the images are 3420 pixels × 2800 pixels to 4077 pixels × 4077 pixels.

(2) Dataset 2: The dataset contains 120 pairs of UAV images. The images include scenes of mountains, cities, coasts, and so on. These images have rigid or affine transformations. The size of the images are 4096 pixels × 3240 pixels and 4320 pixels × 4320 pixels.

(3) Dataset 3: The dataset contains 120 pairs of satellite images. The images include scenes of mountains, cities, farmland, deserts, and so on. The images have rigid transformation. The size of the images are 5120 pixels × 5120 pixels.

To ensure the effectiveness of the proposed algorithm, two methods are used to calculate the parameters: (1) 60% of the images are randomly selected as the training set to calculate the parameters of the proposed algorithm, and the remaining 40% of the images are used as the testing set. (2) 60% of the images are randomly selected as the training and validation, and use 5-fold cross-validation to calculate the parameters of the proposed algorithm. While the remaining 40% of the images are used as test set.

### B. FEATURE POINT EXTRACTION WITHOUT VALIDATION SET

#### 1) THE NUMBER OF FEATURE POINTS EXTRACTED

It is necessary to ensure that the stitching image has good quality and efficiency for the number of feature points extracted. Therefore, the following four aspects would be analyzed: the matching rate of feature points, the descriptors calculate region ratio, SSIM [35] and PSNR. The formula for the ratio of descriptors calculated region is as follows:

$$A_{ratio} = \frac{S(\bigcup_{i=1}^{n} A_i)}{M \times N} \tag{19}$$

where, $M$ and $N$ are the size of the image calculated by the phase correlation algorithm. $A_i$ is the calculation region of each descriptor, $S(\bigcup_{i=1}^{n} X_i)$ is the region of the union of each feature point calculation region. The larger the value, the higher the utilization ratio of image information.

Setting different number of feature points that are extracted, the changes of these parameters are calculated, and the results are shown in figure 6. It can be seen from figure 6 that the average matching rate of dataset 1, 2 and 3 increases with the increase of the number of feature points. The growth rate increased significantly first and then slowed down. The ratio of descriptor calculated region increases with the increase number of feature points. The growth trend slows down when the number reaches 750. The overall trend is to increase first and then fluctuate for SSIM and PSNR. Four indicators grow slowly after 750, which indicates that the number of feature points contributes little to image stitching after reaching 750. Therefore, the number of feature points should be set above 750. Observe the average value of the three data sets of SSIM and PSNR. Both SSIM and PSNR reach the maximum when the number of feature points reaches 900. And the average matching rate and the ratio of descriptor calculated region perform well when the number of feature points is 900. In addition, setting a larger number ensures good results on different datasets due to 900 significantly greater than 750. So, the number of feature points extracted in this paper is set to 900.

#### 2) FEATURE POINT EXTRACTION RATIO

The number of each layer of each octave is calculated according to formula (17) after the total number of feature points is obtained. However, $\alpha$ and $\beta$ are uncertain in formula (17) thus the different $\alpha$ and $\beta$ impact is calculated by the setting training to get values of $\alpha$ and $\beta$. As shown in figure 7, the effect on the average match rate is only showed since the change of $\alpha$ and $\beta$ has no significant effect on SSIM, PSNR and the ratio of descriptor calculated region. It can be also seen from the figure 7 that the overall trend is to increase first and then decrease with the increase of $\beta$. The average matching rate reaches the maximum when $\beta$ in the range of 1.00-1.02. It shows that the feature points of each layer have the same contribution to the final matching in the same octave in DOG. Although the average matching rate reaches the maximum value when $\beta$ =1.01, the value of $\beta$ is set to 1 considering that each layer in the same octave of DOG is equally likely to provide effective feature points. The maximum value of the average matching rate corresponds to $\alpha$ is 0.24 when $\beta$ is 1 in the figure. While in the whole figure, the average matching rate reaches the maximum when $\alpha$ is 0.25 and $\beta$ is 1.01. The image needs to be down-sampling when establishing the Gaussian pyramid, which results in the size of the next octave of images being 1/4 of the previous octave. The value of $\alpha$ is set to 0.25 since the ratio of the number of feature points that can be extracted should also be 1/4.

#### 3) FEATURE POINT EXTRACTION RANGE

Since the large size of the high-resolution image, the range of extreme point comparison is small and the stability is poor when the traditional SIFT algorithm is used to detect extreme points. It may lead to too many feature points gathering in regions with complex texture, which is not conducive to the use of image information. Detecting extreme points from a wider range will be more conducive to the stability of feature points and the use of image information. While a wider range will also lead to greater computational complexity. The impact of different detection ranges on the matching rate and time is analyzed to obtain a low time cost and an appropriate detection range. The results are shown in figure 8. The average matching rate is the lowest when the extreme point detection range is $3 \times 3 \times 3$ (traditional SIFT detection range), indicating that the stability of the feature points is poor at this time. The larger range of feature point detection, the more information will be involved in the calculation. The points with small differences are deleted as well as the extracted feature points can maintain the differences in a larger range, which makes the feature points more stable. The average matching rate shows an overall trend of increasing with the increase of detection range, which indicates that the stability of feature points increases with the increase of detection range. However, the overall growth trend slows down when the detection range is $7 \times 7 \times 3$. The average matching rate grows slowly in the range of $7 \times 7 \times 3$ to $11 \times 11 \times 3$, indicating that the increase of the detection range at this time has little effect on the average matching rate. After the detection range is $11 \times 11 \times 3$, the overall growth trend is larger, and the average matching rate increases faster. The time keep growing exponentially with the increase of detection range. The time is relatively small when the detection range is $7 \times 7 \times 3$. While the time is relatively large when the detection range is after $11 \times 11 \times 3$. It is not conducive to the overall efficiency of the algorithm. The detection range is set to $7 \times 7 \times 3$ considering the time and average matching rate. The feature points extracted from the range of $7 \times 7 \times 3$ is more stable than those from the range of $3 \times 3 \times 3$ (traditional SIFT detection range), however they also incur more time overhead.

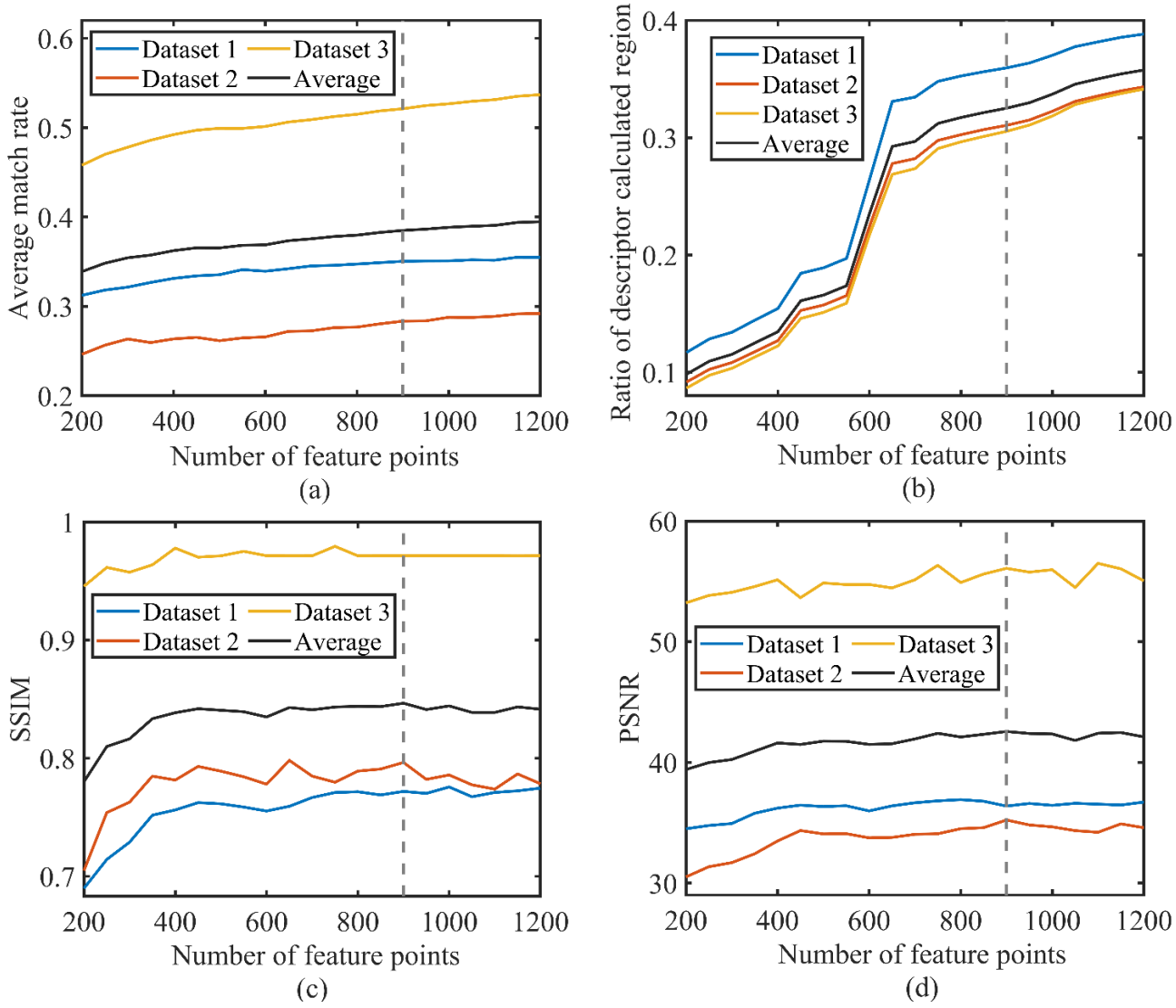


**FIGURE 6.** Influence of the number of feature points on the stitching result: (a) Average match rate; (b) Ratio of descriptor calculated region; (c) SSIM; (d) PSNR.

#### 4) RANGE OF NMS

The effects of different ranges are analyzed to determine the range of NMS. As shown in figure 9, both the average number of feature points extracted and matches decrease with the increase of the NMS range. While the average number of matches decreases slowly between 0 and 0.005. The average matching rate and the ratio of descriptor calculation region increased first and then decreased. And the maximum value is 0.011. A few matched feature points are not conducive to the calculation of the projection transformation matrix of the two images. As much as possible, more matched feature points should be guaranteed. Therefore, priority should be given to the number of extracted feature points and the average matched number compared with the average matching rate and the ratio of descriptor calculation region. The average matching rate and the ratio of descriptor calculation region reach the maximum value when the NMS range ratio is 0.011. However, the number of extracted feature points and the average matched number are poor that above value is not considered. When the range ratio of NMS is 0.005, the average matching rate and the ratio of descriptor calculation region do not reach the maximum value, but it is significantly improved compared with the method without NMS. The number of feature points that can be extracted and matched is relatively large, while the very small difference between methods without NMS. We set the range ratio of NMS to 0.005 considering that the number of extracted feature points and the number of feature point matching are more important. At this time, the number of feature points and the number of feature point matched perform better. The average matching rate and the ratio of descriptor calculation region are significantly improved compared with the method without NMS.

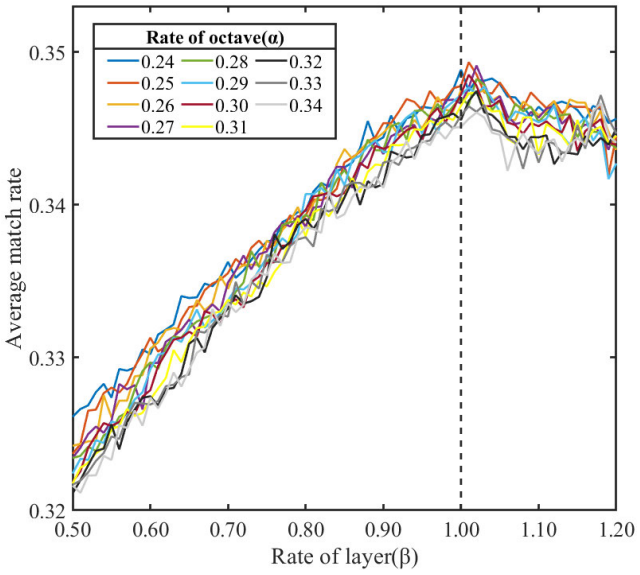


**FIGURE 7.** Effect of feature point extraction ratio on final matching results.

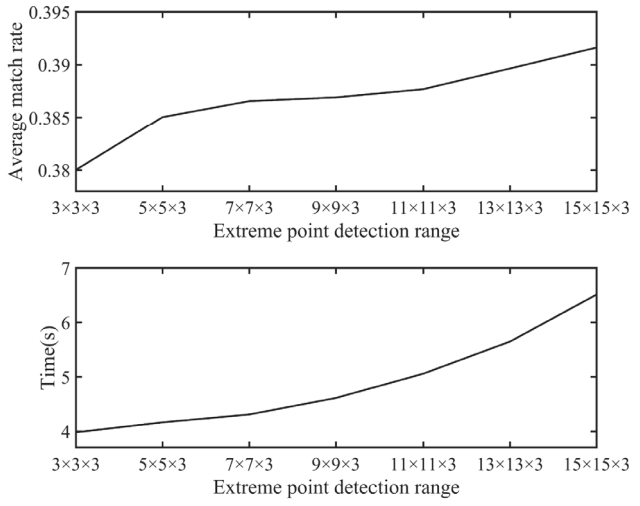


**FIGURE 8.** Influence of detection range of extreme points for feature points.

### C. FEATURE POINT DESCRIPTOR WITHOUT VALIDATION SET

#### 1) PERIPHERAL RING DIVISION AND INNER RING RADIUS RATIO

It is analyzed that the influence of the ratio of $R_1$ to $R_2$ and the division of peripheral rings on the matching rate of feature points through training set. The results are shown in figure 10. The average matching ratio is relatively low when the number of peripheral rings is 3, 4 and 5. The matching rate is relatively high and close when the peripheral ring division is greater than or equal to 6. Therefore, the number of peripheral rings is set to 6 in this paper, which makes the generated descriptor dimension smaller and has a good matching rate. Only observing the line is 6 with the ring division. It can be found that the average matching first increases and then decreases with the increase of the inner ring radius. The maximum value is achieved when the ratio of the inner ring radius is 0.3. Therefore, the ratio of the inner ring radius is set to 0.3.

#### 2) DESCRIPTOR DOWN-SAMPLING

Four down-sampling methods are analyzed for average matching rate and time based on the above set parameter to verify the validity of the method. The four down-sampling methods are: ① The method proposed in this paper (figure 5); ② The method of down-sampling (figure 5(b)) is directly used without parity detection; ③ The sampling interval of down-sampling is changed from 1 to 2 based on ①; ④ The sampling interval of down-sampling is changed from 1 to 2 on the basis of ② The results are shown in figure 11. The average matching rate first increases and then decreases with the increase of the corresponding scale multiple of the feature points. The average matching rates of down-sampling methods ① and ② are close to those of without down-sampling. The average matching rates of down-sampling methods ③ and ④ decrease greatly. It indicates that the description of feature points is incomplete with a larger sampling interval, resulting in a lower matching rate of feature points. When the corresponding scale multiple of the feature points is 10. The average matching rate reaches the maximum in the methods without using down-sampling, methods ① and ② The average matching rate of method ① is higher than that of method ②, which shows that the proposed de sampling method is effective. As shown in figure 11(b), the time increases with the corresponding scale multiple of the feature points. The gap between the four down-sampling methods and the without down-sampling method increases gradually. The larger the sampling interval, the fewer data to be calculated and the smaller the time cost. In addition, the time of sampling method ① is slightly longer than that of sampling method ②. And that of sampling method ③ is longer than that of sampling method ④. Above results are caused since mainly the need of parity detection to select different down-sampling methods. To sum up, the corresponding scale multiple of the feature points is set to 10 when calculating the descriptor. The proposed down-sampling method guarantees a high matching rate with less time overhead. Therefore, the proposed method is effective.

### D. FEATURE POINT EXTRACTION WITH VALIDATION SET

#### 1) THE NUMBER OF FEATURE POINTS EXTRACTED

Setting different number of feature points that are extracted, the changes of these parameters are calculated. The results on the training set are shown in figure 12, and the results on the validation set are shown in figure 13. It can be seen from figure 12 that the average matching rate of 5-fold cross-validation increases with the number of feature points. The growth rate increased significantly first and then slowed down. The ratio of descriptor calculation region for 5-fold cross-validation increases with the number of feature points.

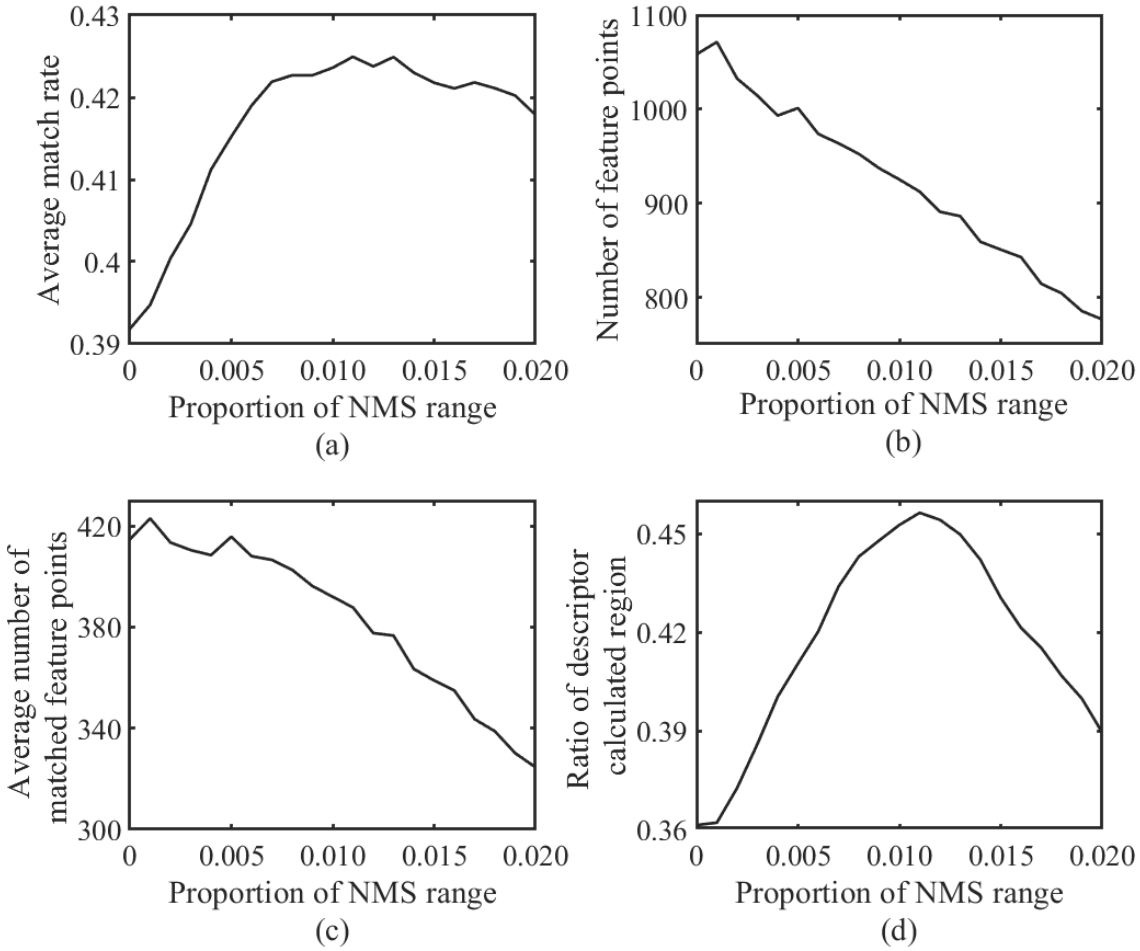


**FIGURE 9.** Influence of the range of NMS on feature point extraction: (a) Average match rate; (b) Number of feature points; (c)Average number of matched feature points; (d) Ratio of descriptor calculated region.

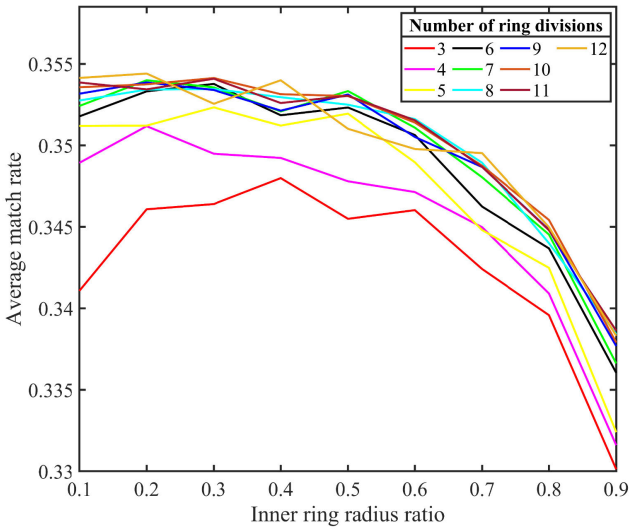


**FIGURE 10.** Influence of peripheral ring division and inner ring radius ratio on feature points.

The growth trend slows down when this number reaches 750. The overall trend is to increase first and then fluctuate for SSIM and PSNR. SSIM and PSNR achieve good results when the number of feature points is 750 and 900 in the validation set, but 900 is better than 750 in the average matching rate and the ratio of the descriptor calculation region. Depending on the performance of the validation set, the number of feature points can be set to 900. Figure 13 shows that the overall variation of the four parameters is similar to that of the training set. When the number of feature points is 900, four parameters obtained good results, and the maximum values of SSIM and PSNR are reached. Therefore, the four parameters on the training and validation sets perform well when the number of feature points is set to 900.

### 2) FEATURE POINT EXTRACTION RATIO

The results of the 5-fold cross-validation for $\alpha$ and $\beta$ are shown in figure 14. It can be seen from figure 14 that the average matching rate of the 5-fold cross-validation reached the maximum when $\beta$ is 1.00-1.02 in the training set. The value of $\beta$ is set to 1 considering that each layer in the same octave of DOG is equally likely to provide effective feature points. In the validation set, the average match rate achieved good results when $\beta$ is1. When $\beta$ is 1.00-1.02, the average matching rate reaches the maximum in three cross-validation

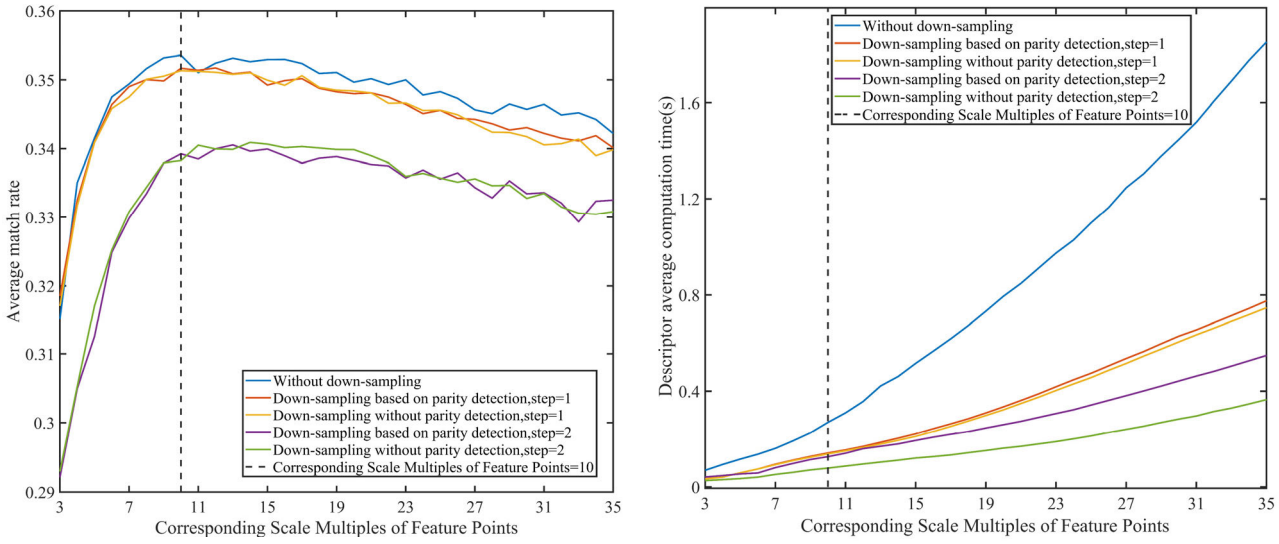


**FIGURE 11.** Comparison of different down-sampling methods: (a) Average match rate; (b) Time.

experiments. When $\beta$ is 1.16 and 1.17 respectively, the average matching rate reaches the maximum value in the two cross-validation experiments. Therefore, the value of $\beta$ Set to 1, it performs well in the training set and validation set. In the training set, when $\beta$ is 1, the maximum of the average match rate corresponds to $\alpha$ is 0.24. Considering that the size of the next octave is 1/4 of the previous octave in the DOG, the value of $\alpha$ is set to 0.25. In the validation set, when $\alpha$ is 0.24, the average matching rate achieved good results. Therefore, the value of $\alpha$ set to 0.25, it performs well in the training set and validation set.

#### 3) FEATURE POINT EXTRACTION RANGE

The influence of different detection ranges on the matching rate and time is analyzed through the 5-fold cross validation, and the results are shown in figure 15. In the training set, the average matching rate is the lowest when the extreme point detection range is $3 \times 3 \times 3$. The average matching rate generally shows a trend of increasing with the detection range, which indicates that the stability of the feature points increases with the detection range. However, the overall growth trend slowed down when the average match rate is $5 \times 5 \times 3$. It shows that the increase of detection range has little effect on the average matching rate at this time. Therefore, the detection range can be set to $5\times 5 \times 3$ based on the training set. In the validation set, the overall growth trend also slowed down when the average match rate is $5 \times 5 \times 3$. In terms of time, the time increases exponentially with the increase of detection distance in training set and test set. When the detection range is $5\times 5 \times 3$, the average matching rate achieves good performance in the validation set, and the time cost is small. Therefore, set the detection range to $5 \times 5 \times 3$. At this time, both training set and validation set can achieve good results. From $5 \times 5 \times 3$ The feature points extracted by the range are more stable than those extracted by the traditional SIFT algorithm, but they also incur more time overhead.

#### 4) RANGE OF NMS

The effects of different ranges are analyzed using a 5-fold cross-validation, as shown in figure 16 and figure 17. In the training set, the results of the 5-fold cross-validations are similar to the results of the method without validation set, so we set the range ratio of NMS to 0.005 based on the training set. In the validation set, the number of feature points and average matched feature points perform well when the NMS is set to 0.005. In terms of the average matching rate and the ratio of descriptor calculation region, better results are obtained than those without NMS. So set the NMS range ratio to 0.005.

### *E. FEATURE POINT DESCRIPTOR WITH VALIDATION SET*

#### 1) PERIPHERAL RING DIVISION AND INNER RING RADIUS RATIO

As shown in figure 18, the ratio of $R_1$ to $R_2$ and the influence of the division of the periphery ring on the matching rate of feature points are analyzed through 5-fold cross-validation. In the five cross-validations of the training set, the average matching rate is relatively low when the number of peripheral rings is 3, 4, and 5. The matching rate is relatively high and close when the peripheral ring division is greater than or equal to 6. Observation line 6 shows that the maximum value is reached when the ratio of inner ring radius is 0.2 and 0.3. In the five experiments, 0.2 appeared once and 0.3 appeared four times. The average value of the five experiments is

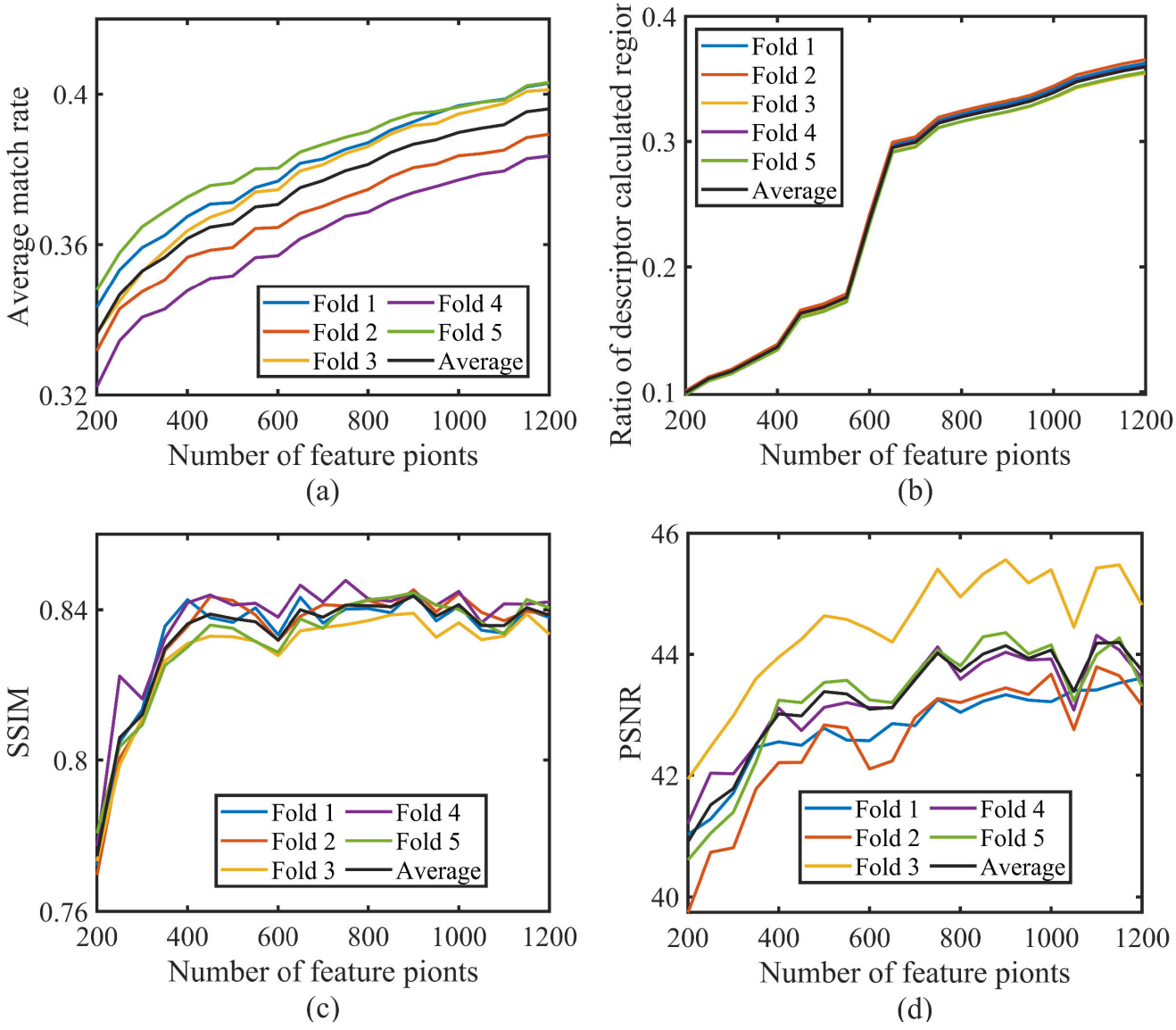


**FIGURE 12.** The influence of the number of feature points in the training set: Average match rate; (b) Ratio of descriptor calculated region; (c) SSIM; (d) PSNR.

close to the results of the five experiments. Based on the training set, the number of peripheral ring and the radius ratio of inner rings can initially set to 6 and 0.3. In the five cross-validation experiments of the validation set, the average matching rate is close when the number of peripheral rings is greater than or equal to 6. Therefore, it is reasonable to set the number of peripheral rings to 6. Observation line 6, the maximum value is reached when the radius ratio of inner ring is 0.2, 0.3, 0.4 and 0.5. In the five experiments, 0.2 appeared twice, and the rest appeared once. From the average value of five cross-validation experiments, the result of the inner ring radius ratio of 0.3 is slightly better than that of 0.2. It shows that the number of peripheral rings is 6 and the radius ratio is 0.3, which can achieve good results, and its setting is reasonable.

### 2) DESCRIPTOR DOWN-SAMPLING

The method without validation set have shown that the method of step=2 is ineffective, so only the method of step=1 is analyzed in 5-fold cross validation. In addition, the experiments without validation set shows the change of time cost, and the results of a five-fold cross-validation are the same as those of previous experiments, so the time overhead in this section is not analyzed. The results of the descriptor down-sampling method are shown in figure 19. Figures 19 (a) and (b) show that in the training set, the maximum values are reached when the corresponding scale multiple of the feature points are 10 and 12 in method ① of 5-fold cross-validation experiment. The effect of 12 is better than the effect of 10 combined with the average values of five experiments. Based on the training set, the corresponding

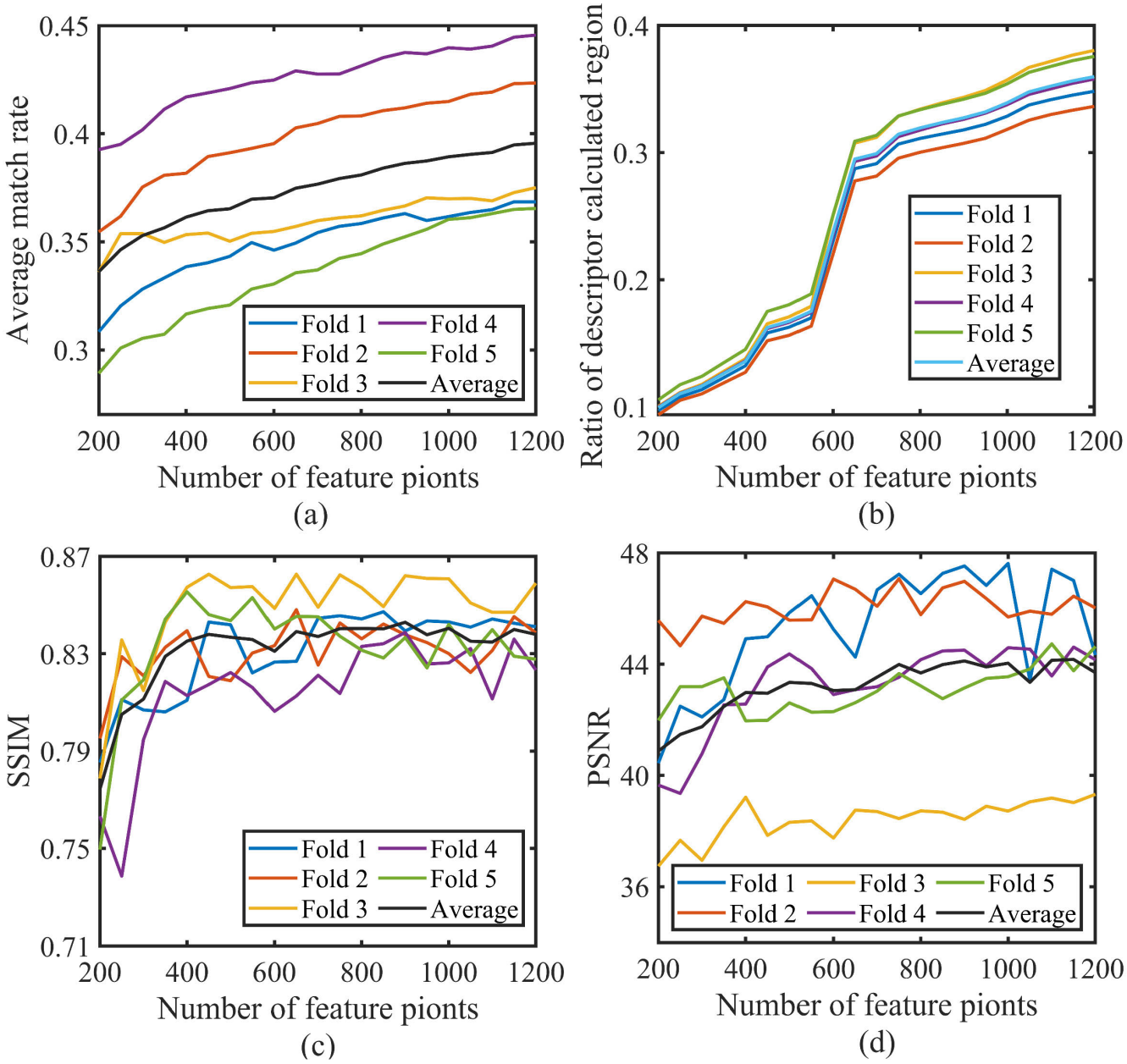


**FIGURE 13.** The influence of the number of feature points in the validation set: (a) Average match rate; (b) Ratio of descriptor calculated region; (c) SSIM; (d) PSNR.

scale multiple of the feature points is set to 12. In the validation set, when the corresponding scale multiple of the feature points is 12, the average matching rate achieved good resultant, and the average value of five experiments reached the maximum value. Therefore, it is reasonable that set the corresponding scale multiple of feature points to 12. Figures 19 (c) and (d) show that method ② achieves good results when the corresponding scale multiple of the feature points is 10 in the training set and validation set. As shown in figures 19 (e) and (f), the maximum value of method ① is greater than the maximum value of method ②, which shows that the parity validation method class of method ① effectively improves the average matching rate. In addition, when the corresponding scale multiple of the feature points is 12, the average matching rate of method ① is close to that of method of without down-sampling, but the time overhead of method ① is smaller, which shows that method ① has a good effect and can effectively reduce the time overhead. Therefore, the method ① will be used and the corresponding scale multiple of the feature points will be set to 12.

### F. ANALYSIS OF EXTRACTION METHODS AND DESCRIPTORS

To test the effect of the feature points extracted by the proposed algorithm, the feature points are extracted on the

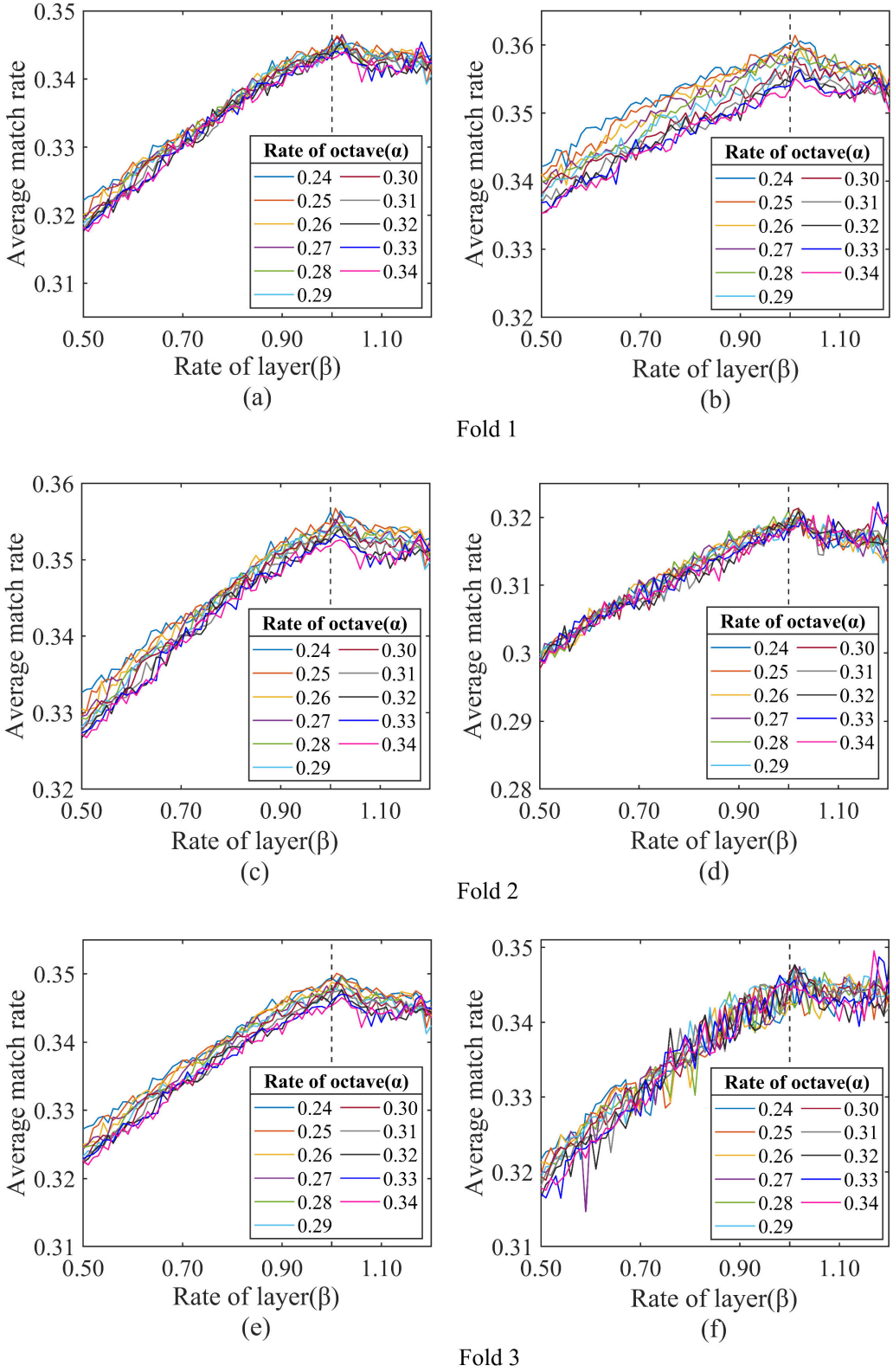


**FIGURE 14.** Effect of feature point extraction ratio on final matching results: (a), (c), (e), (g), (i) and (k): training set; (b), (d), (f), (h) (j) and (l): validation set.

testing set using the proposed algorithm without validation set, the proposed algorithm with validation set and SIFT, respectively. And the feature points are matched by descriptor of SIFT. The results are shown in table 1. The average value of the parameters used in the proposed algorithm without validation set is better than that of the SIFT algorithm, and the

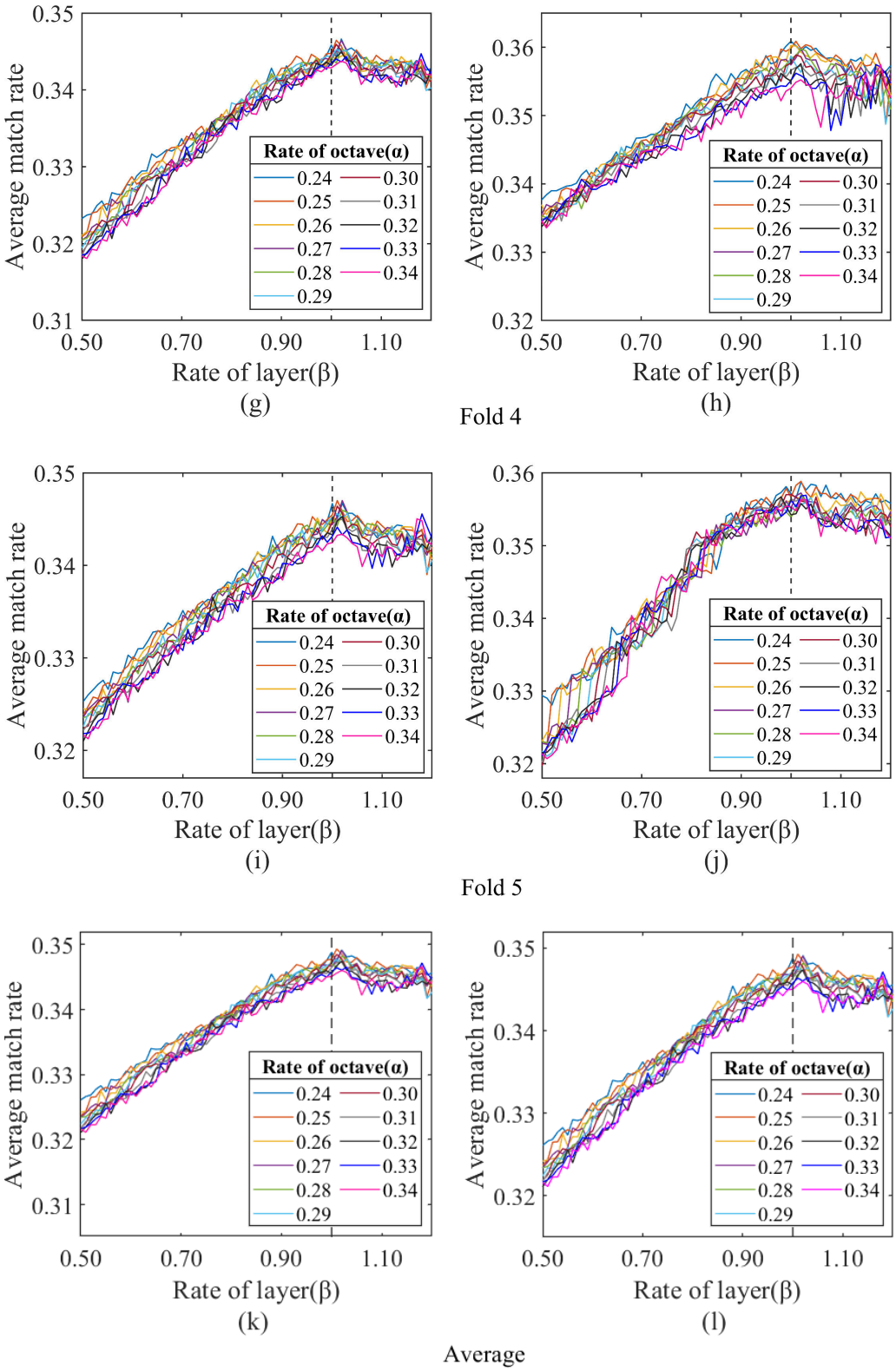


**FIGURE 14.** *(Continued.)* **Effect of feature point extraction ratio on final matching results: (a), (c), (e), (g), (i) and (k): training set; (b), (d), (f), (h) (j) and (l): validation set.**

variance of each aspect is very small. Therefore, the proposed algorithm can extract feature points better than the SIFT algorithm in the case without validation set. In the case with validation set, the average value of parameters used in the proposed algorithm is better than that of the SIFT algorithm, and the variance of each aspect is very small. Compared with

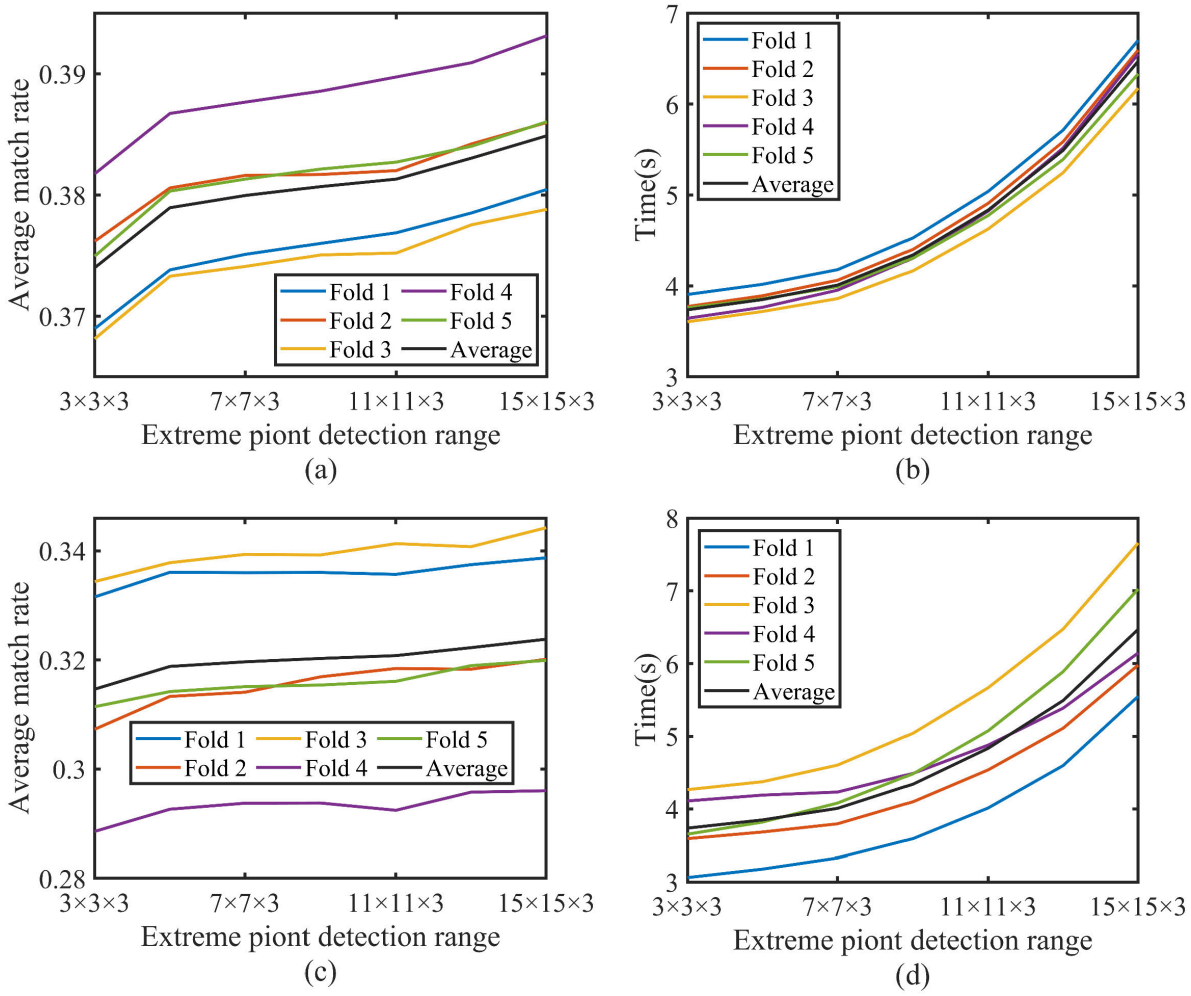


**FIGURE 15.** Influence of detection range of extreme points for feature points: (a) and (b): training set; (c) and (d): validation set.

**TABLE 1.** Comparison of feature point extraction methods.

| Algorithm | Ratio of descriptor calculation region | | Rough matching rate | | PSNR | | SSIM | |
|---|---|---|---|---|---|---|---|---|
| | Average | Std.Dev | Average | Std.Dev | Average | Std.Dev | Average | Std.Dev |
| SIFT | 0.3367 | 0.0941 | 0.3964 | 0.2690 | 33.4373 | 16.7694 | 0.8817 | 0.2807 |
| Proposed algorithm without validation set | **0.3915** | **0.0901** | 0.4989 | 0.2350 | 33.7110 | 16.4891 | 0.8845 | **0.2488** |
| Proposed algorithm with validation set | 0.3913 | 0.0904 | **0.5000** | **0.2328** | **33.7404** | **16.1530** | **0.8855** | 0.2496 |

the case without validation set, the parameters calculated with validation set have better effect in rough matching rate, PSNR and SSIM. Therefore, parameters calculated with validation set can extract feature points better than those calculated without validation set.

Likewise, to test the effectiveness of the descriptor of the proposed algorithm, the feature points are extracted on the testing set using the proposed method. And then feature points are also matched using the descriptor of the proposed algorithm and the descriptor of SIFT, respectively. As shown in table 2, in the case without validation set, the average value of descriptor of SIFT is larger in terms of rough matching rate. However, the average value of the descriptor of the proposed algorithm is superior to the SIFT descriptor in mismatch rate, PSNR and SSIM. It indicates that the descriptor of the proposed algorithm has better performance,

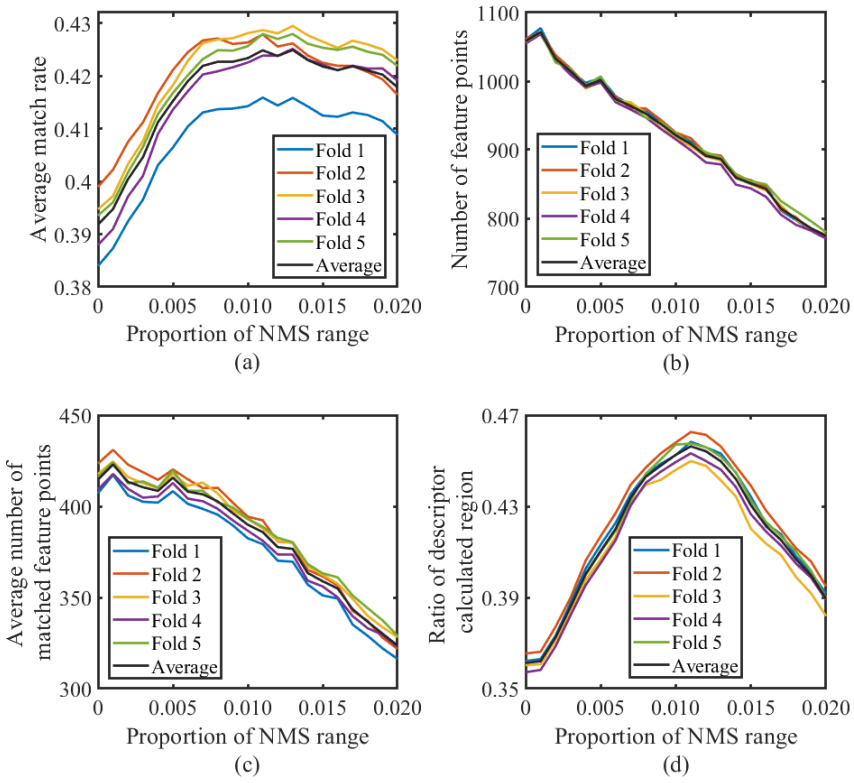


**FIGURE 16.** Influence of the range of NMS in the training set: (a) Average match rate; (b) Number of feature points; (c)Average number of matched feature points; (d) Ratio of descriptor calculated region.

**TABLE 2.** Comparison of descriptors.

| Algorithm | | Rough matching rate | | Mismatch Rate | | PSNR | | SSIM | |
|---|---|---|---|---|---|---|---|---|---|
| | | Average | Std.Dev | Average | Std.Dev | Average | Std.Dev | Average | Std.Dev |
| Proposed method without validation set | SIFT | 0.4989 | 0.2350 | 0.2976 | 0.3588 | 33.7110 | 16.4891 | 0.8845 | 0.2488 |
| | Proposed | 0.4961 | **0.2318** | 0.2837 | 0.3484 | **34.2229** | 15.7710 | **0.9026** | 0.2353 |
| Proposed method with validation set | SIFT | **0.5000** | 0.2328 | 0.2967 | 0.3570 | 33.7404 | 16.1530 | 0.8855 | 0.2496 |
| | Proposed | 0.4953 | 0.2321 | **0.2801** | **0.3419** | 34.1364 | **15.6579** | 0.8949 | **0.2106** |

which can effectively remove mismatches in rough matching. The variance of each aspect is very small. This shows that the fluctuation in all aspects is very small. Therefore, the descriptor of the proposed algorithm has better performance than the descriptor of SIFT. In the case with validation set, the results are similar to those without validation set. In terms of rough matching rate, the case without validation set has better performance due to the larger calculation range of the case with validation set, which can better remove mismatches. The average value and variance of the case without validation set are greater than those of the case with validation set in terms of mismatch rate, PSNR and SSIM. It indicates that the comprehensive performance without validation set is better, but the relative fluctuation is large.

### G. TIME OF IMAGE STITCHING

Three pairs of representative images in the testing set are selected for display to verify the time improvement of the proposed algorithm through reference [26], [28], [36], SIFT algorithm, SURF algorithm, the proposed algorithm with validation set and the proposed algorithm without validation set. The result of rough matching is shown in figure 20. The size of the first pair of images is 5120 pixels × 5120 pixels. The size of the second pair of images is 3500 pixels × 3500 pixels. And the size of the third pair of images is 3800 pixels × 3456 pixels.

Figure 20 shows that the proposed algorithm has the minimum number of incorrect matches, which indicates that the proposed algorithm can effectively extract and describe

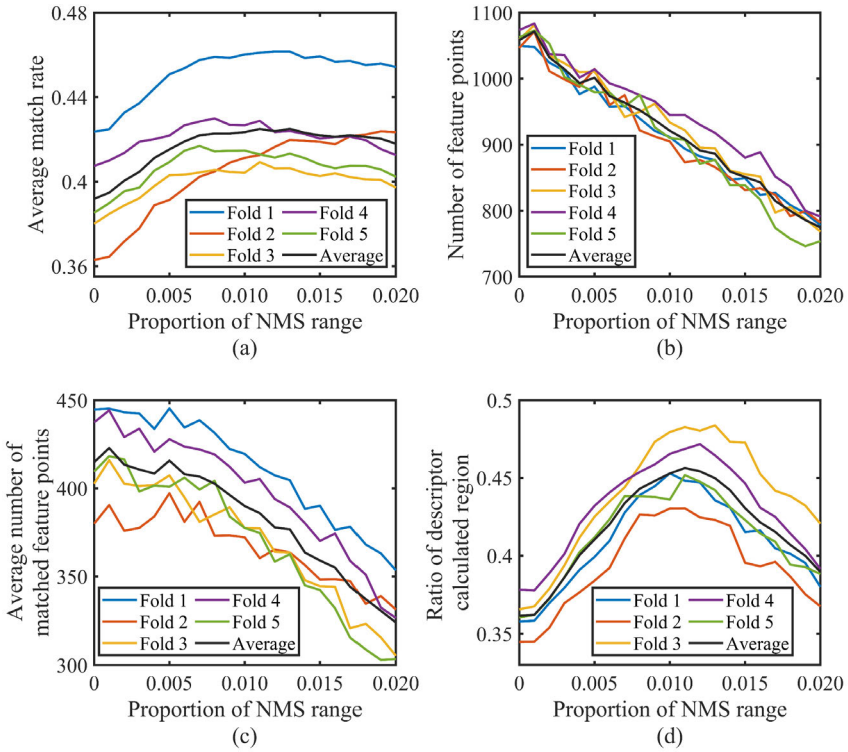


**FIGURE 17.** Influence of the range of NMS in the validation set: (a) Average match rate; (b) Number of feature points; (c)Average number of matched feature points; (d) Ratio of descriptor calculated region.

feature points. The method of reference [28] has a relatively small number of incorrect matches, while the remaining algorithms have a huge number of incorrect matches. The result of the correct match is obtained by the RANSAC algorithm, which is shown in figure 21. As shown in figure 21, the number of feature points matched by references [26], [28], and [36] and SIFT algorithm is too large, which indicates that the number of feature points detected by this algorithm is very large. The total number of feature points extracted is significantly reduced by the proposed algorithm. Although the number of feature point matches of the proposed algorithm is also significantly reduced, the proposed algorithm has a higher correct match rate than other algorithms.

For the 3 pairs of images in figure 20, the time cost is calculated using the five algorithms mentioned above, and the result is shown in table 3. Reference [36] applied Census and NCC algorithms in the feature point matching stage. Therefore, there is not the time of descriptor generation.

As can be seen from table 3, the traditional SIFT algorithm detects many feature points, which results in a large amount of computation and time overhead in each phase of the SIFT algorithm. The method of reference [26] has a relatively small number of feature points compared with the traditional SIFT algorithm. Although the time cost of generating the descriptor is significantly higher, the generated descriptor dimension is smaller. The time cost of feature point matching phase is significantly reduced. And the total time is relatively lower. Reference [28] shows a total time is relatively low due to the relatively low time cost in descriptor generation and feature matching. The method in reference [36] shows fewer feature points and the fast overall speed relatively. However, the total time is still large. The SURF algorithm has a few feature points, and the matching phase takes less time because the descriptor dimension is small. The overall time of the SURF algorithm is relatively small.

The proposed algorithm reduces the computation of the non-overlapping region of the image through the phase correlation algorithm, thus avoids unnecessary time. For the feature point detection, although the NMS and the range of extreme point detection is increased, the number of feature points is limited by algorithm. Therefore, relatively small-time decrease the time cost in the subsequent descriptor generation and matching stages. The proposed descriptor dimension is lower in descriptor generation and feature matching, which further reduces the time overhead of the matching phase. In the case without validation set, compared with the other five algorithms, the proposed algorithm in each stage achieved the least time, which is 2-3 orders of magnitude faster than the SIFT algorithm. The proposed algorithm also achieved good results in the case with validation set. The total time of the proposed algorithm in the two cases is close, accounting for only 0.13%, 0.3% and 0.18% of the SIFT algorithm. Compared with the case without validation set, the case with validation set takes less time to detect

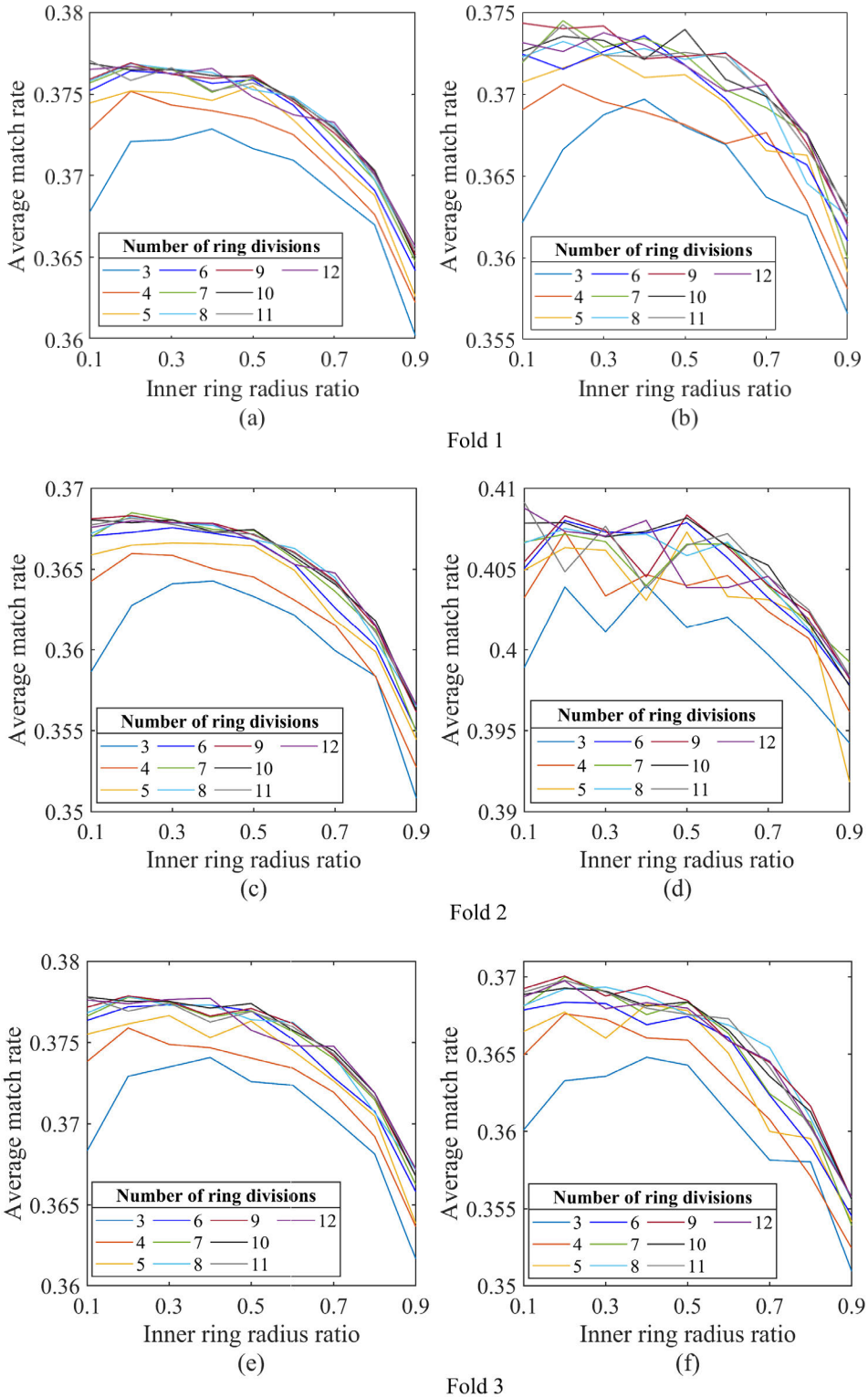


**FIGURE 18.** Influence of peripheral ring division and inner ring radius ratio on feature points: (a), (c), (e), (g), (i) and (k): training set; (b), (d), (f), (h), (j) and (l): validation set.

feature points. More time is spent to generate descriptors due to the smaller detection range of feature points in the case of validation set, while the calculation range of descriptors is larger. Therefore, in both cases, the proposed algorithm greatly reduces the time cost of high-resolution image stitching.

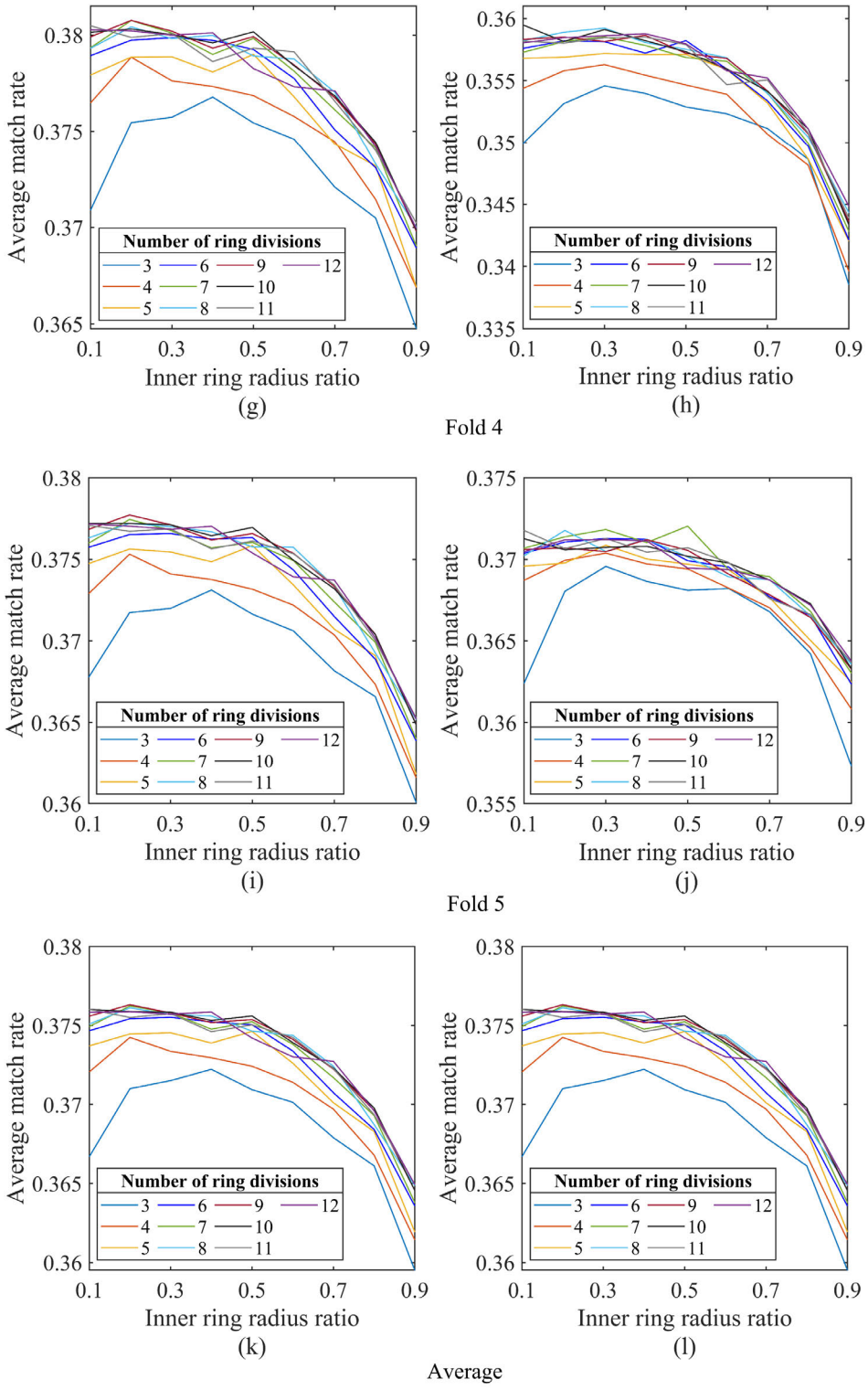


**FIGURE 18.** *(Continued.)* **Influence of peripheral ring division and inner ring radius ratio on feature points: (a), (c), (e), (g), (i) and (k): training set; (b), (d), (f), (h), (j) and (l): validation set.**

### *H. IMAGE STITCHING QUALITY*

The stitching results of the three pairs of images in figure 22 and figure 23. It can be seen from figure 22 that all algorithms have completed image stitching. As shown in figure 23, the method of reference [36] has some stitching traces in image pair 1. The remaining algorithms have no stitching traces

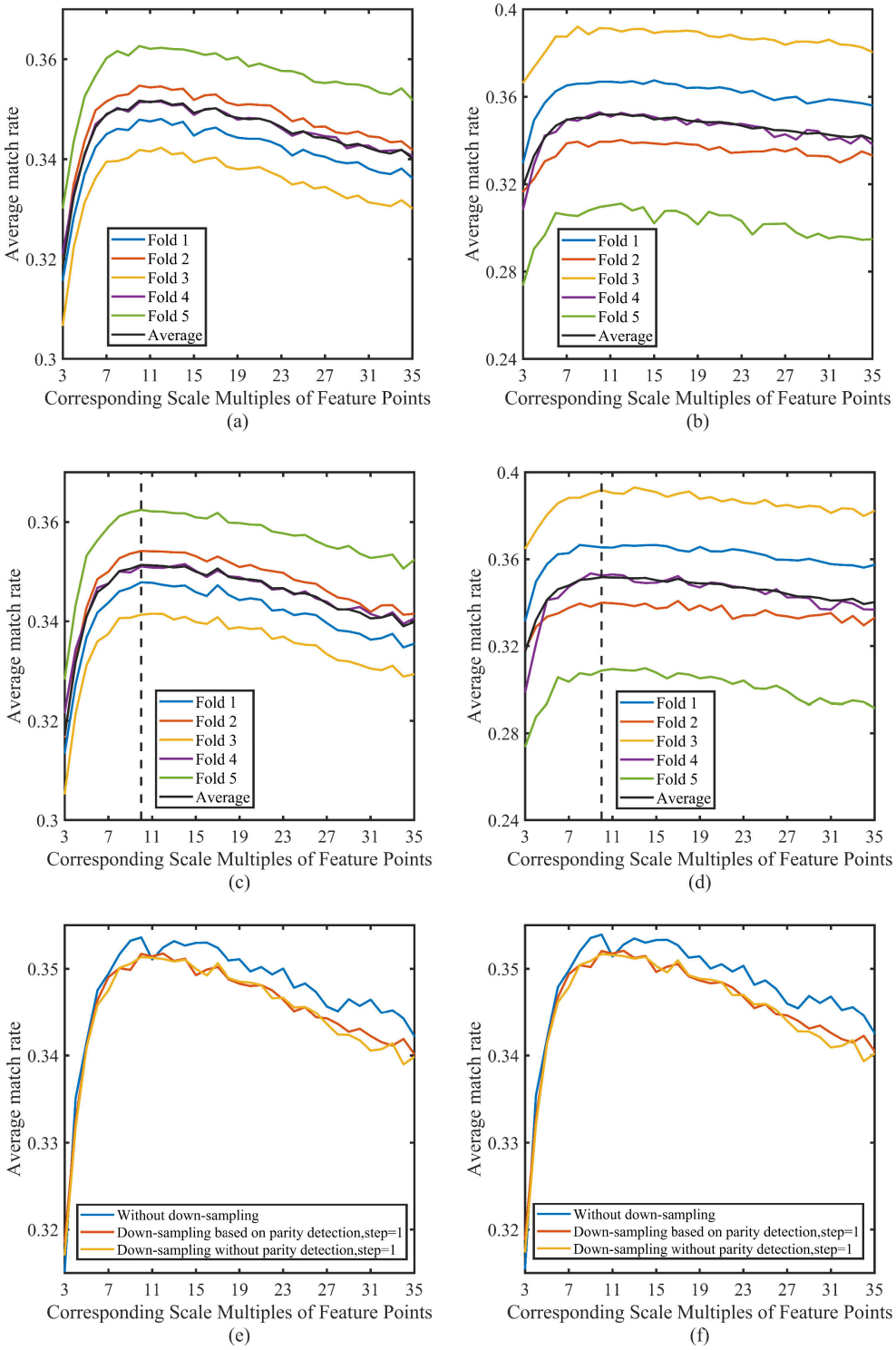


**FIGURE 19.** Influence of different down-sampling methods: (a): down-sampling based on parity detection in training set; (b): down-sampling based on parity detection in validation set; (c): down-sampling without parity detection in training set; (d): down-sampling without parity detection in validation set; (e): Comparison of average values of 5-fold cross-validation in training set; (f): Comparison of average values of 5-fold cross-validation in validation set.

and achieve good visual stitching results. Additionally, the rough matching rate, mismatch rate, SSIM and PSNR of three pairs of image stitching are calculated in figure 20 to further verify the effect of the proposed algorithm on image stitching

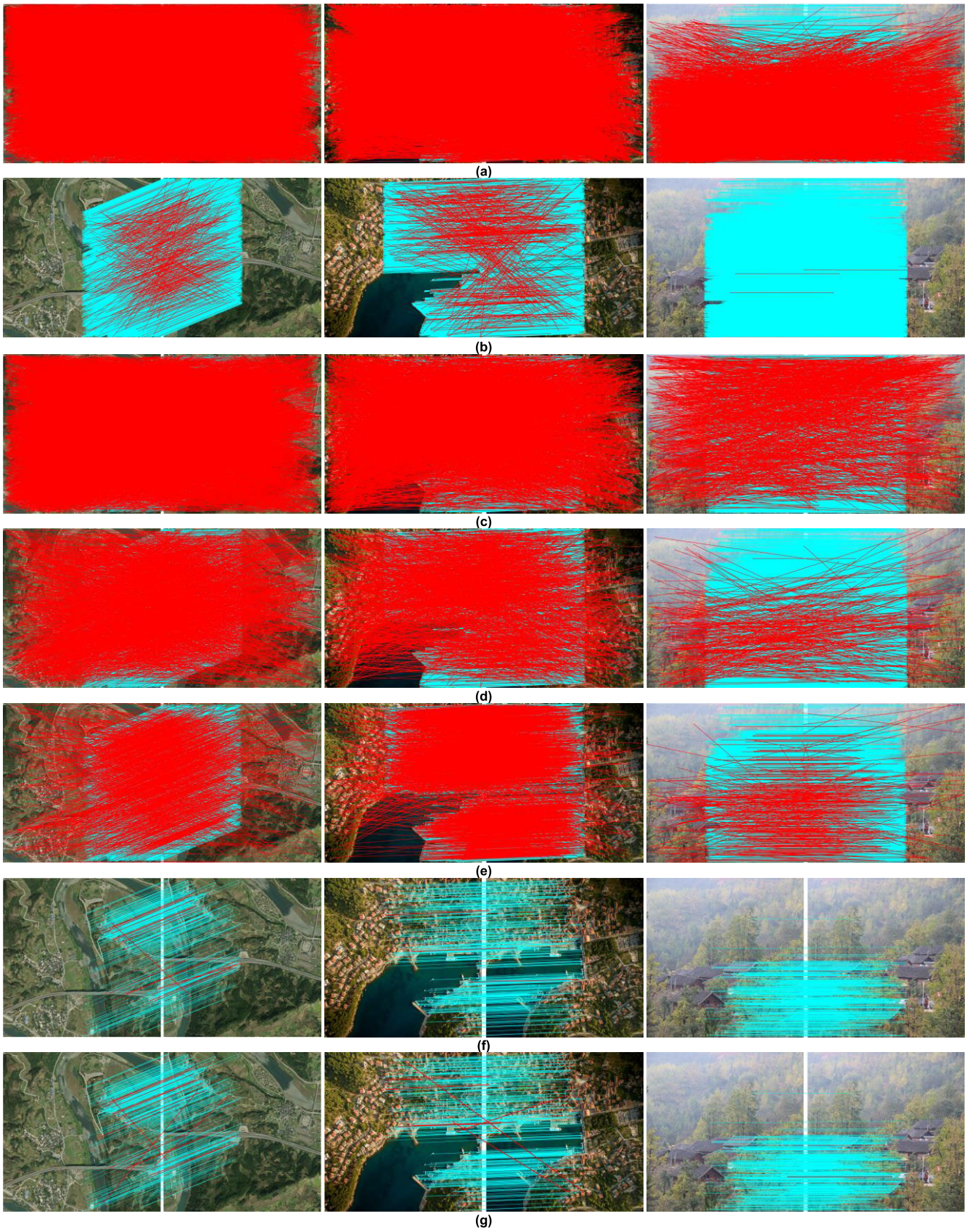


**FIGURE 20.** Result of rough matching (The red line indicates incorrect matches, while the blue line indicates correct matches): (a) Reference [26]; (b) Reference [28]; (c) Reference [36]; (d)SIFT; (e)SURF; (f) Proposed algorithm without validation set; (g) Proposed algorithm with validation set.

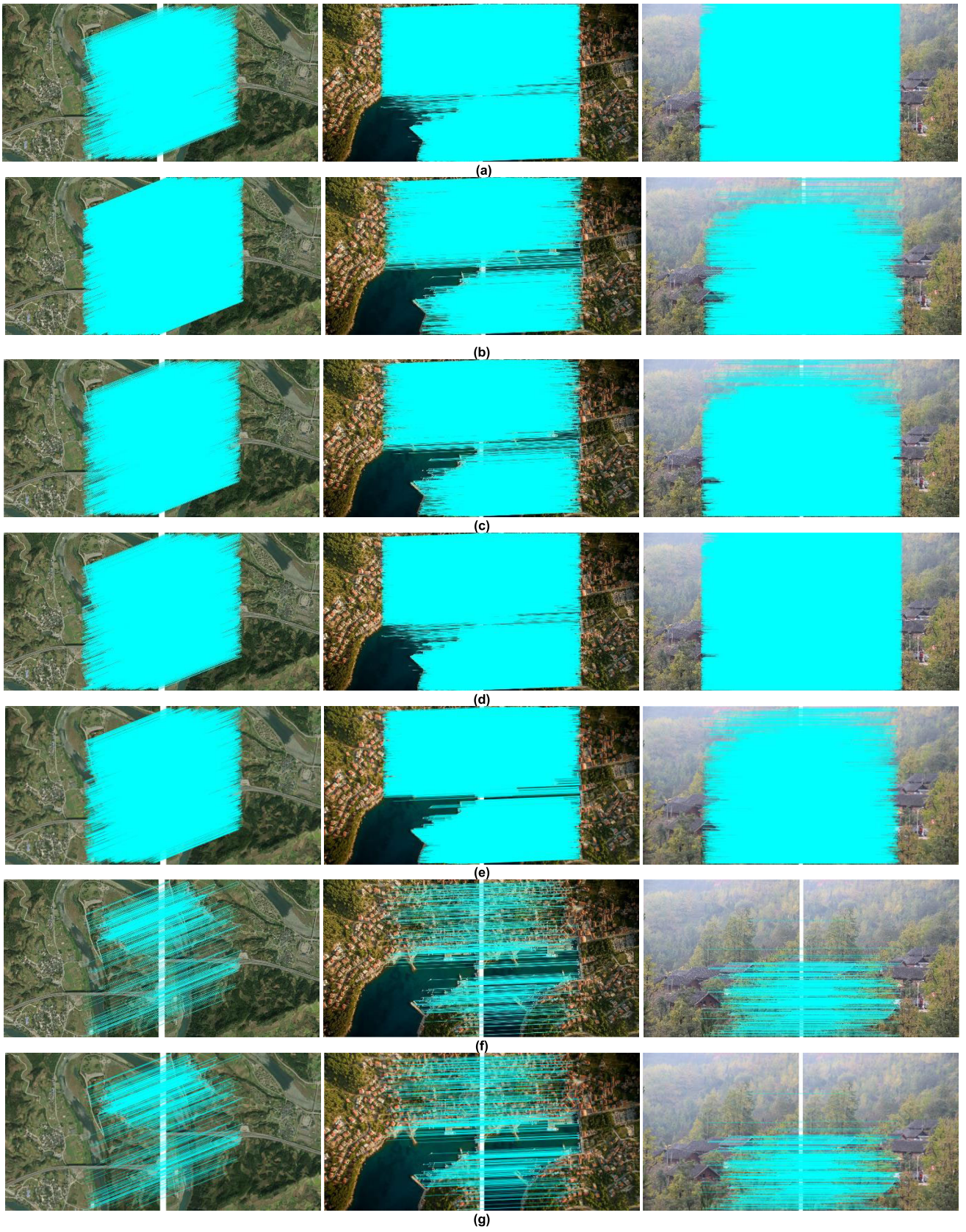


**FIGURE 21.** Correctly matched results: (a) Reference [26]; (b) Reference [28]; (c) Reference [36]; (d)SIFT; (e)SURF; (f) Proposed algorithm without validation set; (g) Proposed algorithm with validation set.

**TABLE 3. Comparison of image alignment times.**

| Images | Algorithm | Number of feature points | Time of phase correlation algorithm (s) | Time of feature point detection (s) | Time of descriptor generation (s) | Time of feature point matching (s) | Total time (s) |
|---|---|---|---|---|---|---|---|
| Image pair 1 | Reference [26] | 136147,149953 | 1.34 | 32.72 | 192.74 | 1501.84 | 1727.30 |
| | Reference [28] | 63022,62888 | | 11.09 | 13.74 | 1006.88 | 1033.05 |
| | Reference [36] | 55788,57830 | | 15.61 | | 3050.20 | 3065.81 |
| | SIFT | 176704, 193103 | | 11.55 | 26.69 | 9855.13 | 9893.37 |
| | SURF | 60338，64016 | | 42.38 | 41.70 | 538.16 | 622.24 |
| | Proposed algorithm without validation set | 1003,1003 | | **6.13** | **0.38** | 5.08 | **12.93** |
| | Proposed algorithm with validation set | 1003,1004 | | **6.13** | 0.52 | **5.06** | 13.05 |
| Image pair 2 | Reference [26] | 62642,55273 | 0.66 | 13.10 | 29.76 | 251.90 | 294.76 |
| | Reference [28] | 41916,41937 | | 8.73 | 4.61 | 362.64 | 375.98 |
| | Reference [36] | 22743,22765 | | 5.73 | | 546.85 | 570.57 |
| | SIFT | 103620,104105 | | 14.32 | 10.88 | 2625.22 | 2650.42 |
| | SURF | 41493,37116 | | 13.48 | 28.48 | 211.89 | 253.85 |
| | Proposed algorithm without validation set | 1111,1117 | | 4.51 | **0.45** | **2.38** | 8.00 |
| | Proposed algorithm with validation set | 1111,1117 | | **4.28** | 0.54 | 2.40 | **7.88** |
| Image pair 3 | Reference [26] | 68582,77995 | 0.72 | 13.34 | 36.64 | 378.66 | 428.64 |
| | Reference [28] | 55538,55495 | | 13.02 | 10.44 | 1064.26 | 1088.44 |
| | Reference [36] | 31318,31791 | | 6.57 | | 941.93 | 948.5 |
| | SIFT | 144612,160476 | | 18.28 | 14.92 | 4763 | 4796.2 |
| | SURF | 35889,39745 | | 19.70 | 28.16 | 210.42 | 258.28 |
| | Proposed algorithm without validation set | 1025,1026 | | 5.03 | **0.37** | 2.61 | 8.73 |
| | Proposed algorithm with validation set | 1025,1026 | | **4.87** | 0.46 | **2.56** | **8.61** |

quality. The results are shown in table 4. The reference [28] achieved the largest value in the rough matching rate of feature points, which indicates that the utilization rate of feature points detected is better. In both cases, the proposed method is only lower than that in reference [28]. And the utilization rate of feature points detected is obviously superior to the other four algorithms, which shows that it has good robustness in the extracted feature points. The overall mismatch rate of the proposed algorithm is very low. The proposed algorithm has the lowest mismatch rate without a validation set. This indicates that the reliability is excellent applying the new descriptor. Reference [36] shows the best results in the aspects of SSM and PSNR. There is no significant difference between the proposed algorithm, reference [26] and [36], SURF and SIFT. The quality of image stitching is improved less when feature points are too many. The proposed algorithm has a good stitching effect.

### *I. ALGORITHM PERFORMANCE TEST*

Moreover, five algorithms are tested on testing set to analyze the effectiveness of the algorithm and the results are shown in tables 5-7. In tables 5-7, mismatched images do not participate in subsequent parameter calculations. Reference [36] shows the best effect in terms of SSIM and PSNR. While the stitching effect of new algorithm is slightly less than that of SIFT algorithm. In the absence of validation set, the proposed algorithm, reference [28] and SIFT have the best performance for the number of mismatched image pairs. For mismatch rate, the new algorithm has the best performance in test sets 1 and 2 of datasets. Reference [28] has the best performance in test set of dataset 3, however, it has little gap with the new algorithm. This shows, the new algorithm has the best performance combining the results of above three datasets. The new algorithm achieves an excellent performance in consideration of the data of the number of mismatched image pairs and mismatch rate. It shows that the new algorithm can extract feature points effectively. The extracted feature points have good stability. And the proposed descriptor also has good description ability. The time cost of the new algorithm is significantly better than the other five algorithms. The average time (9.60s, 13.46s and 15.81s) is 0.86%, 0.43% and 0.10% respectively of the SIFT algorithm. The overall

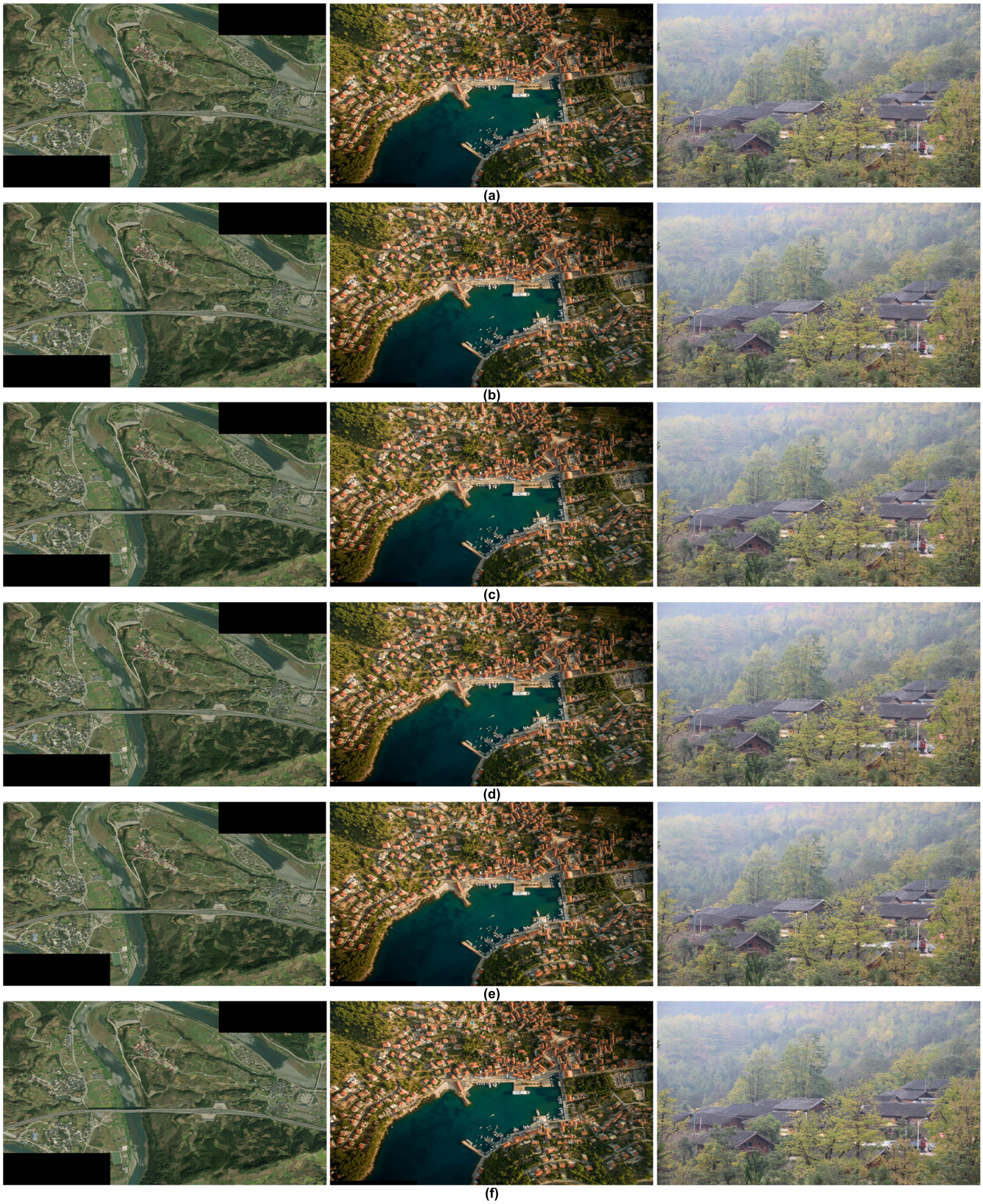


**FIGURE 22.** Results of image stitching: (a) Reference [26]; (b) Reference [28]; (c) Reference [36]; (d)SIFT; (e)SURF; (f) Proposed algorithm without validation set; (g) Proposed algorithm with validation set.

(g)

**FIGURE 22.** *(Continued.)* **Results of image stitching: (a) Reference [26]; (b) Reference [28]; (c) Reference [36]; (d)SIFT; (e)SURF; (f) Proposed algorithm without validation set; (g) Proposed algorithm with validation set.**

**TABLE 4. Comparison of image stitching quality.**

| Images | Algorithm | Rough matching rate | Mismatch rate | SSIM | PSNR |
|---|---|---|---|---|---|
| Image pair 1 | Reference [26] | 35.81%,32.51% | 32.71% | 0.9968 | 40.2294 |
| | Reference [28] | 58.38%,58.51% | 1.90% | 0.9986 | 40.2294 |
| | Reference [36] | 29.43%,28.39% | 58.36% | **0.9995** | **44.6623** |
| | SIFT | 36.18%,33.11% | 13.94% | 0.9986 | 40.2308 |
| | SURF | 30.26%,28.52% | 11.13% | 0.9986 | 40.2288 |
| | Proposed algorithm without validation set | 66.50%,66.50% | **1.37%** | 0.9986 | 40.2342 |
| | Proposed algorithm with validation set | **66.60%,66.53%** | 2.99% | 0.9986 | 40.2350 |
| Image pair 2 | Reference [26] | 39.44%,44.69% | 15.57% | 0.9981 | 36.3424 |
| | Reference [28] | **73.84%,73.81%** | 4.85% | 0.9981 | 36.3427 |
| | Reference [36] | 39.08%,39.03% | 29.06% | **0.9995** | **40.4470** |
| | SIFT | 44.70%,44.49% | 11.49% | 0.9981 | 36.3537 |
| | SURF | 34.33%,38.36% | 19.78% | 0.9981 | 36.3998 |
| | Proposed algorithm without validation set | 55.00%,54.70% | **0.65%** | 0.9981 | 36.3459 |
| | Proposed algorithm with validation set | 55.26%,54.96% | 1.14% | 0.9981 | 36.3451 |
| Image pair 3 | Reference [26] | 76.40%,67.18% | 12.84% | 0.9994 | 44.4505 |
| | Reference [28] | **99.30%,99.37%** | 0.007% | 0.9994 | 44.4505 |
| | Reference [36] | 57.87%,57.01% | 5.22% | **0.9998** | **48.4485** |
| | SIFT | 66.88%,60.27% | 2.54% | 0.9994 | 44.4505 |
| | SURF | 66.36%,59.93% | 1.9% | 0.9994 | 44.4501 |
| | Proposed algorithm without validation set | 74.73%,74.66% | **0%** | 0.9994 | 44.4500 |
| | Proposed algorithm with validation set | 74.82%,74.75% | 1.3% | 0.9994 | 44.4499 |


**FIGURE 23.** Details of stitched images: (a) Stitched image detail regions (The red box regions); (b) Reference [26]; (c) Reference [28]; (d) Reference [36]; (e)SIFT; (f)SURF; (g) Proposed algorithm without validation set; (h) Proposed algorithm with validation set.

speed of new algorithm reaches 2-3 orders of magnitude faster than that of SIFT algorithm. Meanwhile, the smallest standard deviation is obtained applying new algorithm, which indicates that our algorithms presented excellent performance on different images.

Compared with the method without validation set, the method with validation set is better in terms of SSIM and PSNR, but worse in terms of mismatch rate and time. The method with validation set has a small detection range of feature points, resulting in less stable extracted feature points and

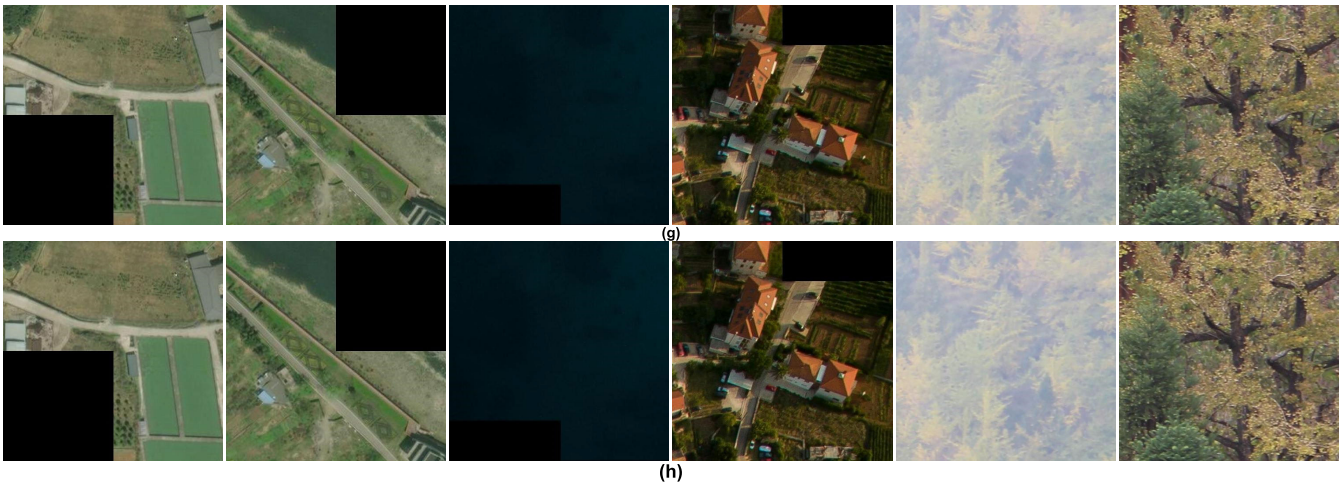

**FIGURE 23.** *(Continued.)* **Details of stitched images: (a) Stitched image detail regions (The red box regions); (b) Reference [26]; (c) Reference [28]; (d) Reference [36]; (e)SIFT; (f)SURF; (g) Proposed algorithm without validation set; (h) Proposed algorithm with validation set.**

**TABLE 5. Algorithm comparison on dataset 1.**

| Algorithm | Number of mismatched Image pairs | Mismatch rate | | SSIM | | PSNR | | Times(s) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Average | Std.Dev | Average | Std.Dev | Average | Std.Dev | Average | Std.Dev |
| Reference [26] | 20 | 0.3131 | 0.2093 | **0.9782** | **0.0602** | **40.4252** | 13.2867 | 167.02 | 231.93 |
| Reference [28] | 8 | 0.1277 | 0.2071 | 0.9459 | 0.1026 | 36.1885 | 14.7913 | 240.21 | 409.75 |
| Reference [36] | 7 | 0.2493 | 0.2275 | 0.9638 | 0.0895 | 40.0933 | 15.0634 | 281.68 | 205.61 |
| SIFT | 6 | 0.2211 | 0.2334 | 0.9449 | 0.1035 | 36.0014 | 15.0277 | 1110.38 | 1320.58 |
| SURF | 9 | 0.2331 | 0.2546 | 0.9544 | 0.0826 | 36.9501 | 15.0996 | 117.42 | 129.37 |
| Proposed algorithm without validation set | **4** | **0.0982** | **0.1730** | 0.9566 | 0.0824 | 36.1878 | 12.8835 | **9.64** | **2.26** |
| Proposed algorithm with validation set | 5 | 0.1471 | 0.2101 | 0.9624 | 0.0700 | 36.5313 | **12.2755** | 14.41 | 4.59 |

**TABLE 6. Algorithm comparison on dataset 2.**

| Algorithm | Number of mismatched Image pairs | Mismatch rate | | SSIM | | PSNR | | Times(s) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Average | Std.Dev | Average | Std.Dev | Average | Std.Dev | Average | Std.Dev |
| Reference [26] | 11 | 0.4646 | **0.3319** | 0.8949 | 0.1160 | 30.2809 | **14.9259** | 198.70 | 184.12 |
| Reference [28] | **3** | 0.4327 | 0.3820 | 0.8539 | 0.1594 | 29.7258 | 15.7530 | 830.05 | 1160.46 |
| Reference [36] | 9 | 0.4796 | 0.3848 | **0.9519** | **0.0633** | **35.7036** | 15.0437 | 3122.22 | 2896.11 |
| SIFT | 6 | 0.4595 | 0.3824 | 0.8823 | 0.1300 | 30.8722 | 15.7065 | 3151.13 | 2932.14 |
| SURF | 6 | 0.4834 | 0.3915 | 0.8824 | 0.1281 | 30.6260 | 15.0506 | 250.95 | 237.29 |
| Proposed algorithm without validation set | 7 | **0.3401** | 0.3350 | 0.8796 | 0.1307 | 30.8895 | 14.9287 | **13.46** | **2.97** |
| Proposed algorithm with validation set | 9 | 0.3858 | 0.3813 | 0.8919 | 0.1224 | 31.1904 | 15.2400 | 18.81 | 5.64 |

higher mismatch rate. The method of validation set calculates descriptors over a larger range, resulting in greater time costs.

To sum up, the method without validation set has better performance in terms of time cost and mismatch rate, while the method with validation set has better performance in terms of SSIM and PSNR. Both methods have achieved fast speed and good splicing quality. This shows that the new method is feasible, which has certain application value in the field of high-resolution image accelerated stitching.

**TABLE 7. Algorithm comparison on dataset 3.**

| Algorithm | Number of mismatched Image pairs | Mismatch rate | | SSIM | | PSNR | | Times(s) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Average | Std.Dev | Average | Std.Dev | Average | Std.Dev | Average | Std.Dev |
| Reference [26] | 1 | 0.3867 | 0.1643 | 0.9768 | 0.0049 | 42.1722 | 12.6077 | 767.40 | 782.60 |
| Reference [28] | 2 | **0.0426** | **0.0428** | 0.9968 | 0.0050 | 43.3331 | 11.6035 | 1255.05 | 1664.70 |
| Reference [36] | **0** | 0.1186 | 0.1083 | **0.9991** | **0.0009** | **48.6951** | 11.6276 | 2728.34 | 941.26 |
| SIFT | 1 | 0.1163 | 0.0481 | 0.9970 | 0.0049 | 42.6568 | **11.1931** | 15212.28 | 6465.62 |
| SURF | **0** | 0.0835 | 0.0481 | 0.9971 | 0.0049 | 42.6216 | 12.2533 | 394.16 | 439.83 |
| Proposed algorithm without validation set | 2 | 0.0599 | 0.0972 | 0.9970 | 0.0049 | 42.5688 | 12.5432 | **15.81** | **2.27** |
| Proposed algorithm with validation set | 2 | 0.1289 | 0.1745 | 0.9971 | 0.0049 | 42.6776 | 12.5191 | 22.63 | 3.25 |

## V. CONCLUSION

In conclusion, a new fast stitching algorithm for high-resolution image is presented based on SIFT. The overlap region of images is computed by phase correlation algorithm to reduce the non-overlapping region calculations. The more stable and better spatially distributed obtained when expand detection range of extreme points and add NMS. Meanwhile, the range of the descriptor is calculated by the method of down-sampling, and the circular descriptor is constructed with only 56-dimensional. This indicates that the total descriptor calculation is faster in lower dimensions by the new algorithm. Importantly, the new algorithm shows the greatly decreased time cost and relatively good results in image stitching quality since the average time of calculations (9.60s, 13.46s and 15.81s) is 2-3 orders of magnitude faster than the SIFT algorithm. A good stitching quality and excellent performance are achieved for high-resolution image applying the new algorithm. It is also proved that our algorithm has the potential application value in real-time image stitching, especially decrease of time cost. However, the result of the proposed algorithm in image stitching quality is unsatisfactory, especially when processing UAV images data. Therefore, in further work, we will research the improvement of image stitching quality. On the one hand, we will try to extract more stable feature points. On the other hand, we will try to improve the descriptor's ability to handle more complex transformations in the image.

## REFERENCES

[1] D. Ghosh and N. Kaabouch, "A survey on image mosaicing techniques," *J. Vis. Commun. Image Represent.*, vol. 34, pp. 1–11, Jan. 2016.
[2] W. Ma, Z. Wen, Y. Wu, L. Jiao, M. Gong, Y. Zheng, and L. Liu, "Remote sensing image registration with modified sift and enhanced feature matching," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 1, pp. 3–7, Jan. 2016.
[3] H.-H. Chang, G.-L. Wu, and M.-H. Chiang, "Remote sensing image registration based on modified SIFT and feature slope grouping," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 9, pp. 1363–1367, Sep. 2019.
[4] M. Koeva, R. Bennett, and C. Persello, "Remote sensing for land administration 2.0," *Remote Sens.*, vol. 14, no. 17, p. 4359, Sep. 2022.
[5] N. Himthani, M. Brunn, J.-Y. Kim, M. Schulte, A. Mang, and G. Biros, "CLAIRE—Parallelized diffeomorphic image registration for large-scale biomedical imaging applications," *J. Imag.*, vol. 8, no. 9, p. 251, Sep. 2022.
[6] S. K. Sharma and K. Jain, "Image stitching using AKAZE features," *J. Indian Soc. Remote Sens.*, vol. 48, pp. 1389–1401, Sep. 2020.
[7] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571.
[8] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2548–2555.
[9] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
[10] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, Jun. 2008.
[11] Z. Zhang, L. Wang, W. Zheng, L. Yin, R. Hu, and B. Yang, "Endoscope image mosaic based on pyramid ORB," *Biomed. Signal Process. Control*, vol. 71, Jan. 2022, Art. no. 103261.
[12] J. Xu, "Fast image registration method based on Harris and SIFT algorithm," *Chin. Opt.*, vol. 8, no. 4, pp. 574–581, 2015.
[13] V. S. Bind, P. R. Muduli, and U. C. Pati, "A robust technique for feature-based image mosaicing using image fusion," in *Proc. Int. J. Adv. Comput. Res.*, Mar. 2013, p. 263.
[14] P. Kang and H. Ma, "An automatic airborne image mosaicing method based on the SIFT feature matching," in *Proc. Int. Conf. Multimedia Technol.*, Jul. 2011, pp. 155–159.
[15] A. Laraqui, A. Saaidi, and K. Satori, "MSIP: Multi-scale image preprocessing method applied in image mosaic," *Multimedia Tools Appl.*, vol. 77, no. 6, pp. 7517–7537, Mar. 2018.
[16] M. Yan, D. Qin, G. Zhang, P. Zheng, J. Bai, and L. Ma, "Nighttime image stitching method based on guided filtering enhancement," *Entropy*, vol. 24, no. 9, p. 1267, Sep. 2022.
[17] M. Gong, S. Zhao, L. Jiao, D. Tian, and S. Wang, "A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4328–4338, Jul. 2014.
[18] Y. Guang, S. Wenbang, Z. Xingming, and L. Tongshao, "Accumulated error elimination method of aerial array image stitching," *Infr. Laser Eng.*, vol. 50, pp. 353–361, Sep. 2021.
[19] S. Zhao, G. Yu, and Y. Cui, "New UAV image registration method based on geometric constrained belief propagation," *Multimedia Tools Appl.*, vol. 77, no. 18, pp. 24143–24163, Feb. 2018.
[20] X. Gong, Y. Liu, and Y. Yang, "Robust stepwise correspondence refinement for low-altitude remote sensing image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 10, pp. 1736–1740, Oct. 2021.
[21] Y. H. Zhang, X. Jin, and Z. J. Wang, "A new modified panoramic UAV image stitching model based on the GA-SIFT and adaptive threshold method," *Memetic Comput.*, vol. 9, no. 3, pp. 231–244, Sep. 2017.
[22] J. Zhao, X. Zhang, C. Gao, X. Qiu, Y. Tian, Y. Zhu, and W. Cao, "Rapid mosaicking of unmanned aerial vehicle (UAV) images for crop growth monitoring using the SIFT algorithm," *Remote Sens.*, vol. 11, no. 10, p. 1226, May 2019.
[23] B. Kupfer, N. S. Netanyahu, and I. Shimshoni, "An efficient SIFT-based mode-seeking algorithm for sub-pixel registration of remotely sensed images," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 379–383, Feb. 2015.

[24] J. Ma, J. Jiang, H. Zhou, J. Zhao, and X. Guo, "Guided locality preserving feature matching for remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4435–4447, Aug. 2018.
[25] Q. Zeng, J. Adu, J. Liu, J. Yang, Y. Xu, and M. Gong, "Real-time adaptive visible and infrared image registration based on morphological gradient and C_SIFT," *J. Real-Time Image Process.*, vol. 17, no. 5, pp. 1103–1115, Oct. 2020.
[26] Y. Chen, M. Xu, H.-L. Liu, W.-N. Huang, and J. Xing, "An improved image mosaic based on Canny edge and an 18-dimensional descriptor," *Optik*, vol. 125, no. 17, pp. 4745–4750, Sep. 2014.
[27] H. Shi, L. Guo, S. Tan, G. Li, and J. Sun, "Improved parallax image stitching algorithm based on feature block," *Symmetry*, vol. 11, no. 3, p. 348, Mar. 2019.
[28] Y. Wang, Z. Tang, M. Zhong, Y. Wang, R. Zeng, D. Zhu, and C. Yang, "SIFT image stitching algorithm based on phase correlation and texture classification," *Chin. J. Quantum Electron.*, vol. 37, pp. 650–658, Nov. 2020.
[29] H. Zhou, W. Yi, L. Du, and Y. Qiao, "Convolutional neural network-based dimensionality reduction method for image feature descriptors extracted using scale-invariant feature transform," *Laser Optoelectron. Prog.*, vol. 56, no. 14, 2019, Art. no. 141008.
[30] X. Li, Y. Chen, W. Li, Y. Li, L. Zheng, and S. Wu, "Projection aided digital shearography scanning detection technology," *Infr. Laser Eng.*, vol. 50, pp. 325–332, Aug. 2021.
[31] Q. Du, A. Fan, Y. Ma, F. Fan, J. Huang, and X. Mei, "Infrared and visible image registration based on scale-invariant PIIFD feature and locality preserving matching," *IEEE Access*, vol. 6, pp. 64107–64121, 2018.
[32] Y. Li, W. Qiao, H. Jin, J. Jing, and C. Fan, "Reliable and fast mapping of keypoints on large-size remote sensing images by use of multiresolution and global information," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 9, pp. 1983–1987, Sep. 2015.
[33] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.
[34] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "SAR-SIFT: A SIFT-like algorithm for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 453–466, Jan. 2014.
[35] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[36] Z. Tang, Z. Ding, R. Zeng, Y. Wang, J. Wen, L. Bian, and C. Yang, "Multi-threshold corner detection and region matching algorithm based on texture classification," *IEEE Access*, vol. 7, pp. 128372–128383, 2019.

**ZETIAN TANG** received the M.S. degree in electronics and communication engineering from Guizhou University, China, in 2020. He is currently a Lecturer with Liupanshui Normal University. His research interests include image processing and computer vision.

**ZEMIN ZHANG** received the M.S. degree in physical chemistry from Southwest University, China, in 2018. She is currently a Lecturer with Liupanshui Normal University. Her research interest includes computational chemistry.

**WEI CHEN** received the B.Eng. and M.S. degrees from the Lanzhou University of Technology, Lanzhou, China, in 2017 and 2021, respectively. He is currently pursuing the Ph.D. degree in electrical engineering with Sun Yat-sen University, Shenzhen, China. His research interests include signal processing, cognitive radio, and integrated sensing and communications.

**WENTAO YANG** received the M.S. degree in optical engineering from Chongqing Normal University, China. He is currently a Professor with Liupanshui Normal University. His research interests include image processing and optical system design.