

# A similar feature point matching method for aerial electric power tower images based on a one-ring neighborhood of vertices

Keyu Chen\*, Jufu Guo, Weijun Wang, Bo Lei, Hui Huang, and Lijing Luo

Guizhou Power Grid Co., Ltd. Guiyang Power Supply Bureau, Guiyang 550000, Guizhou, China

Received: 29 August 2024 / Accepted: 13 November 2024

**Abstract.** Due to the susceptibility of images to various factors such as scale changes, imaging conditions, and image noise, traditional feature point matching methods are difficult to achieve ideal matching accuracy, leading to many challenges in the automatic analysis and processing of power tower images. To improve the matching accuracy of similar feature points in aerial images of electric power tower, this study proposes a method for matching similar feature points in aerial images of electric power tower based on a vertex ring neighborhood, addressing the problem of large matching errors caused by factors such as scale and imaging conditions. This method first adopts the feature point extraction technique of electric power tower image based on decomposition and filtering, and constructs the Laplacian pyramid of the original image. Subsequently, filter banks are used to decompose the pyramid image in different directions, extract local extremum points as candidate feature point sets, and merge them to obtain the final feature point set. On this basis, taking into account the spatial relationships and geometric characteristics around the feature points, a texture mapping algorithm based on vertex ring neighborhood and patch classification is applied to construct a vertex ring neighborhood structure and extract geometrically representative electric power tower features. Finally, by using a texture feature point matching method based on the principle of similarity, the similarity between the real-time image and the reference image features is calculated for matching, and combined with the Babbitt coefficient to remove mismatched feature points, accurate matching of similar feature points in aerial electric power tower images is achieved. The experimental results show that this method has a mean square error of less than 0.1 Pixel in matching similar texture feature points of aerial power tower images under various working conditions, significantly improving the matching accuracy and providing an effective tool for automatic analysis and processing of power tower images.

**Keywords:** Electric power tower, Aerial images, Vertex one-ring neighborhood, Similar feature point matching, Laplace pyramid, Filters, Babbitt coefficients.

## 1 Introduction

The matching technology of similar feature points in aerial images of electric power tower plays a crucial role in the operation, maintenance, and monitoring of the power industry [1]. However, this technology often faces many challenges in practical applications. Due to differences in image scale, imaging conditions, and other factors, the feature point matching of aerial electric power tower images often has significant errors, which not only affects the accurate positioning of power facilities, but also may pose potential risks to the safety monitoring of power lines [2–4]. The change in image scale leads to changes in the size and position of feature points in the image, thereby

affecting the accuracy of matching. Differences in imaging conditions, including camera models, lens focal lengths, exposure times, etc., also lead to changes in the expression and distribution patterns of image feature points. In addition, factors such as changes in lighting, viewing angles, and noise and distortion in the image can increase the difficulty of feature point extraction and matching, making the matching results susceptible to interference and resulting in false or missed matches. During the inspection process of power towers, if the matching of image feature points is inaccurate, it will cause inspection personnel to be unable to accurately determine the status of the tower, thereby missing the timely detection and handling of potential safety hazards. In addition, although some existing feature point matching methods can achieve the extraction and matching of image feature points to a certain extent, there

\* Corresponding author: [ckycsg@126.com](mailto:ckycsg@126.com)

are still some shortcomings. Some methods generate feature descriptors with high sub dimensions, which increases the difficulty of storage and matching; Some methods encode image features into fixed length binary strings, resulting in the loss of some detailed information; There are also some methods that mainly describe based on the grayscale differences of images, which cannot fully capture the complex texture and detail information of images. These issues have limited the widespread application of feature point matching technology in the power industry.

In the process of image feature point extraction and matching, some existing feature point matching methods have been studied at a certain level. For example, Sourabh *et al.* [5] studied the Scale-Invariant Feature Transform (SIFT) based optical remote sensing image matching method. By extracting key points in the image and generating feature descriptors with rotation and scale invariance high precision matching is realized in remote sensing images with different viewing angles and scales. However, the feature description sub dimension generated by SIFT algorithm is higher, usually reaching 128 dimensions or higher. This increases the difficulty of storage and matching. De *et al.* [6] studied a real-time aerial image mosaic matching method based on parallel hash. This method extracts features from each aerial image through parallel computing, codes the extracted features, converts high-dimensional feature vectors into low-dimensional hash codes, and quickly finds potential matching pairs by comparing the similarity between different image hash codes. But the hash method usually encodes image features into fixed length binary strings, which may lead to the loss of some details. And if there are serious problems such as blurring, noise or distortion in aerial images, it may lead to inaccurate hash coding, thus affecting the matching effect. Meddeber *et al.* [7] proposed an image feature matching method based on the fast Haar-k nearest neighbor algorithm when studying the photometric and geometric stitching problems of remote sensing images. Using the efficient calculation of Haar features and the k-d tree data structure, they realized fast matching of image features, and had certain robustness to interference factors such as light changes, scale changes, and partial occlusion in the image. Through reasonable feature extraction and matching strategy, it can achieve more stable matching results in complex environments. However, Haar features are mainly described based on the gray difference of the image, which may not fully capture the complex texture and detail information of the image. In practical applications, appropriate feature extraction and matching algorithms need to be selected according to specific needs and scene characteristics to achieve better matching results. Ekanayake and Lei [8] studied the point matching method based on Support Vector Regression (SVR). SVR algorithm can achieve accurate fitting of image feature sample data by finding the optimal hyperplane, so it has high accuracy in point matching tasks. The performance of SVR algorithm is affected by many parameters, such as penalty coefficient, kernel function parameters, etc. The selection of parameters needs to be adjusted and optimized according to specific tasks, otherwise the matching performance may be degraded or unstable. Luo [9] proposed a path planning

method for climbing robots based on 3D CAD models to address the problems of low intelligence and low work efficiency during power tower maintenance and repair. This method extracts geometric information from SAT files, constructs tower skeletons using OpenGL, and implements global path planning, significantly improving intelligence and work efficiency. But this method relies on the 3D CAD model of the power tower, and the accuracy of the model directly affects the accuracy and reliability of path planning.

Based on the shortcomings of existing research, this paper proposes a method for matching similar feature points of aerial electric power tower images based on vertex ring neighborhood to improve the matching accuracy of similar feature points. Firstly, by delving into the local and global information of image feature points, more robust feature descriptors are constructed to improve the accuracy and robustness of matching; Secondly, by utilizing the vertex ring neighborhood structure, the spatial relationships and geometric characteristics around the feature points are fully considered, further enhancing the expressive power of the feature descriptors; Finally, by combining the texture feature point matching method based on the principle of similarity and the Babbitt coefficient to remove mismatches, high-precision matching of power tower images was achieved under different scales and imaging conditions. In summary, the proposed method for matching similar feature points in aerial power tower images based on vertex ring neighborhood aims to construct more robust feature descriptors by deeply mining the local and global information of image feature points, in order to achieve high-precision matching of power tower images under different scales and imaging conditions, and provide strong support for intelligent operation and maintenance in the power industry.

## 2 Aerial electric power tower image similar feature point matching method

### 2.1 Decomposition and filtering based feature point extraction method for aerial electric power tower images

The Laplace decomposition of the aerial electric power tower image will be continuously subband decomposition of its bandpass image, and when the directional filter bank is applied to these bandpass subbands, it can be decomposed into directional subbands, and then the directional information can be effectively obtained [10]. Although the directional filter bank can realize the multi-directional decomposition of the image, but due to its lack of multi-resolution characteristics, the low-frequency component of the image is not well handled, and if the directional filter bank is used directly to decompose the image, the low-frequency information will be unevenly distributed into the various directional subbands [11]. In the electric power tower image, most of the feature information exists in the middle and high frequency regions, so the low frequency part will weaken the ability to capture the directional

information, and the usual processing method is to remove the low frequency component of the image [12]. The most direct way to solve this problem here is to remove the low-pass component.

In order to extract the feature points in different scales and directions of the aerial electric power tower image, the Laplace pyramid of the original image is constructed, the low-frequency components are removed [13], and the Laplace pyramid is decomposed in multiple directions using a multidirectional filter bank to extract the feature points in the subbands of each direction. Laplacian pyramid is a multi-resolution image representation method that constructs a series of image layers of different scales by performing continuous Gaussian smoothing and downsampling operations on the image. Each layer of the image represents the information of the original image at a certain scale. Firstly, starting from the original image, a Gaussian pyramid is constructed through a series of Gaussian smoothing and downsampling operations. Each layer of the Gaussian pyramid is obtained by Gaussian smoothing and downsampling the image from the previous layer. Each layer of the Laplacian pyramid is calculated by subtracting the current layer image from the previous layer image of the Gaussian pyramid. This difference image represents the high-frequency part of the image at the current scale, namely the edges and details of the image. In the Laplacian pyramid, the bottom layer contains most of the detailed information of the image, while the higher layers, the more blurry image becomes, representing more low-frequency information. In practical operation, removing low-frequency components is achieved by simply ignoring the upper layers of the Laplacian pyramid. These upper layers mainly contain low-frequency information of the image, which is not necessary for feature point extraction and may even introduce interference. Therefore, in subsequent processing, only focus on layers that retain high-frequency information. In order to remove low-frequency components, it is chosen to only retain the high-frequency part of the Laplacian pyramid and ignore the upper layer containing low-frequency information. In practical applications, based on the characteristics of the image and analysis requirements, the number of reserved layers can be selected to fully reflect the structural information of the power tower. After removing the low-frequency components, perform multi-directional filtering on the retained high-frequency parts. A multi-directional filter bank decomposes an image into multiple directional sub bands, each containing information about the image in that direction. Perform feature point detection in various directional sub bands to extract significant structural or key points. Due to the fact that feature points may have responses in different directions, but the response positions may deviate, it is necessary to merge and optimize the extracted multi-directional candidate feature points. By setting a sliding window, merge the candidate feature points within the window, and determine the final feature point position based on their weights and coordinates. During the filtering process, some feature points only respond significantly in one direction, and cannot be extracted or have weak responses in other sub bands. These points are identified as isolated points. Isolation points are caused by

noise or specific textures in the image, which may not be very helpful for subsequent image analysis or recognition tasks, and are therefore removed during the merging process.

The Laplace pyramid decomposition [14] was performed on the original aerial electric power tower image, and then for the  $w$ th layer image in the pyramid, perform multidirectional decomposition, get the decomposed image on direction  $b$ . Denote the image obtained from the above process as  $A_w^b$ , and then extract the local poles as candidate feature points. In this paper, the extracted feature points of aerial electric power tower image are defined to satisfy:

$$Q_w^b(x, y) = \left\{ (x, y) \left| A_w^b(x, y) \geq \frac{\varpi_1}{v(\varpi, w)} N_w^b(x, y) \right. \right\} \quad (1)$$

$$N_w^b(x, y) = \sum_{x-\frac{w}{2}}^{x+\frac{w}{2}-1} \sum_{y-\frac{w}{2}}^{y+\frac{w}{2}-1} A_w^b(i, j). \quad (2)$$

Among them,  $v(\varpi, w)$  is the size of the feature point extraction window; the  $\varpi_1$  is the lower threshold.  $N_w^b(x, y)$  is the set of candidate feature points extracted in the direction  $b$  of layer  $w$ , the final full set of feature points obtained as  $Q_w^b(x, y)$ ;  $(i, j)$  is pixel pairs.

The feature points in different directions are obtained by filtering, usually a feature point can respond in different directions, due to the feature point extraction in different directions, the position of the same point responding in each direction will have deviation, in order to merge these feature points, it is necessary to extract the set of 8-direction candidate feature points  $N_w^b(x, y)$  perform merge optimization [15].

Set the sliding window of size  $v \times w$ , if the value of the window is too small (*e.g.*,  $3 \times 3$  pixels), the candidate feature points extracted in each direction cannot be completely merged, and if the value of the window is too large (*e.g.*,  $7 \times 7$  pixels), it is easy to contain the isolated noise points extracted, so a sliding window of  $5 \times 5$  pixels can be used to merge the feature points in each direction and reduce the interference of the isolated points.

Suppose there are four candidate feature points in a window of size  $5 \times 5$  pixels, 2 candidate points  $q_{b1} \in N_w^b(x, y)$ ,  $q'_{b1} \in N_w^b(x, y)$ , on the direction  $b = 1$  are extracted, merge the candidate points in the same direction, and get  $q_b \in N_w^b(x, y)$ , determine its weight as 2, then merge the candidate points in other directions and determine the weights, and finally determine the final location of the feature point based on the weights and coordinates of the merged points in each direction [16]. Generally, set the window with  $\varepsilon$  candidate feature points, which are taken from  $\delta$  directions, where in the filtering direction  $b$ , can extract  $\varepsilon_b$  candidate feature points. If  $\varepsilon > \delta$ , combine  $N_w^b(x, y)$  and get its sub-pixel coordinates as:

$$(x_b, y_b) = \left( \frac{\sum x_j}{\varepsilon_b}, \frac{\sum y_j}{\varepsilon_b} \right) (j = 1, 2, \dots, \varepsilon_b). \quad (3)$$

After the candidate feature points in the same direction are merged, the sub-pixel coordinates of the feature points are then determined from the merged points in each direction as:

$$(x_b, y_b) = \left( \frac{\sum x_j x_b}{b}, \frac{\sum y_j y_b}{b} \right). \quad (4)$$

The Laplace transform is performed on the aerial electric power tower image to obtain the high frequency portion. In the process of performing the filtering a portion of the feature points respond only within a certain direction, *i.e.*, when  $\varepsilon = 1$ , the point cannot be extracted in other direction subbands, these points are labeled as isolated points [17]. In image processing, especially when filtering aerial images of power towers, the formation of isolated points is an important phenomenon. Firstly, Laplace transform is applied to aerial images of power towers to extract high-frequency information, such as edges and details, from the images. These high-frequency components contain the structural information of power towers, including their lines, edges, etc. Next, filter the extracted high-frequency part. The filtering here includes directional filtering, bandpass filtering, etc., aiming to further extract specific directional information from the image. In the process of directional filtering, the image is decomposed into multiple directional subbands, each containing information about the image in that direction. Perform feature point detection in various directional subbands. Feature points are prominent structural or key points in an image that have high contrast or specific geometric characteristics. When a feature point has a significant response only in one directional subband and cannot be extracted or has a weak response in other directional subbands, it is judged as an isolated point. The formation of isolated points is due to certain specific structures or texture information in the image being only prominent in one direction, while being suppressed or blurred in other directions. After identifying the isolated points, further processing is carried out to obtain the final feature point set of aerial power tower images. Use a merging process function to integrate feature point information from various directional subbands. This function weights or filters feature points based on their strength, position, direction, and other attributes to obtain a more accurate and stable set of feature points. During the merging process, remove feature points that have been identified as isolated points. Because isolated points are only caused by noise or specific textures in the image, they may not be very helpful for subsequent image analysis or recognition tasks. After merging and removing isolated points, the final feature point set of aerial power tower images was obtained. This feature point set contains significant and stable structural information in the image, which is used for subsequent image matching, recognition, or analysis tasks. Then the final aerial electric power tower image feature point set is determined as:

$$Q_w(x, y) = \bigcap_{b=1}^8 (x_b, y_b) f(Q_w^b(x, y)). \quad (5)$$

Among them,  $f$  is the merge process function.

## 2.2 Texture mapping algorithms based on vertex one-loop neighborhood and facet classification

Considering the set  $Q_w(x, y)$  of image feature points extracted in Section 2.1, the feature points cannot directly

reflect the geometric characteristics of the original image target. In the decomposition and filtering method, the extracted feature points may be unevenly distributed in the image, leading to sparse feature points in some regions, while in other regions, feature points are dense, affecting the accuracy of subsequent feature matching and positioning. One ring neighborhood of a vertex [18] is an important concept in graph theory, computer graphics, computer vision and other fields. It describes the set of vertices directly adjacent to a given vertex in a graph or grid structure, and the connection between these vertices and a given vertex. The texture mapping method based on vertex loop neighborhood and triangular patch classification plays an important role in the matching of similar feature points of aerial electric power tower images. This method takes into account the spatial relationship and geometric characteristics around the feature points, which helps to extract more representative and robust features. Therefore, this paper uses a texture mapping algorithm based on vertex loop neighborhood and patch classification to map and extract texture feature points with spatial and geometric features from the feature point set of aerial electric power tower images for subsequent feature point matching.

The triangular faceted slice classification technique further enhances the accuracy of texture mapping, which can better capture and represent the local details and shape features of the image by texture mapping the feature points of the aerial electric power tower image when the target of the aerial electric power tower image has deformation. This method is especially suitable for aerial electric power tower images, because the structure of the electric power tower often has clear geometric shape and texture features. These geometric and texture features can be more accurately described by triangular facet classification and play a key role in the matching process [19].

In this paper, the initial aerial electric power tower image feature point data is processed before the similar feature point matching the aerial electric power tower image. The main purposes of processing this data are:

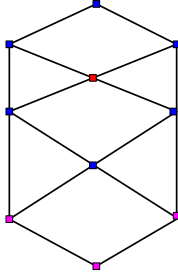
### 2.2.1 Find the one-ring neighborhood of each vertex in the feature points of the aerial electric power tower image

In the set of feature points  $Q_w(x, y)$ , vertex  $Q$  its one-loop neighborhood of the red vertex consists of all the neighboring vertices that are directly connected to it, as shown in Figure 1. The one-ring neighborhood of the red vertex consists of six blue vertices, and the three pink vertices do not belong to the one-ring neighborhood of the red vertex.

### 2.2.2 Classification of triangular facets

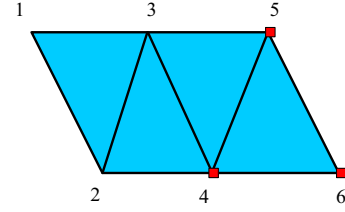
The specific criteria and process for classifying triangular facets are based on their geometric characteristics and their representation in three-dimensional models or images. Generally speaking, the classification of triangular facets is based on their pose, intersection, or other geometric properties. In pose based classification, a fully horizontal triangular patch is a triangular patch with three completely horizontal edges. This type of surface is stable on





**Fig. 1.** Vertex one ring neighborhood.

a horizontal plane and is not easily deformed by gravity. A horizontal triangular surface is a triangular surface with only one side parallel to the horizontal plane. This type of patch has one edge in a suspended state and requires additional support. A regular triangular surface is a triangular surface with three sides that are not parallel to the horizontal plane. These types of patches have more complex geometric shapes and contain more shape features. In the classification based on intersection (assuming intersection with the slice plane), completely non intersecting triangular facets are triangular facets that do not intersect with the slice plane. A triangular surface with intersecting vertices is a triangular surface with only one vertex intersecting the slice plane. A triangular surface with two intersecting vertices is a triangular surface with two vertices intersecting the slice plane. A triangular surface with a vertex and an edge intersecting is a triangular surface with a vertex and an edge intersecting the slice plane. A triangular surface with two intersecting edges is a triangular surface with two edges intersecting the slice plane. A triangular surface with three intersecting vertices is a triangular surface with all three vertices intersecting the slice plane. In the classification based on the direction of the normal vector (for 3D models), the horizontally downward triangular surface is parallel to the normal vector and  $Z$ -axis, and the surface is in an absolute suspended state. The downward sloping triangular surface is at a certain angle between the normal vector and the  $Z$ -axis. When the angle is less than a certain threshold, a support structure needs to be added. By classifying, different types of triangular patches can be more accurately identified, thereby capturing and representing local details of the image more precisely. In the power tower image, the complete horizontal triangular patch represents the bottom or platform part of the tower, while the inclined triangular patch represents the slant support or inclined part of the tower. In 3D modeling, the construction process of the model is optimized through classification. For triangular facets that require the addition of supporting structures, they should be identified and processed in advance during the modeling process to avoid deformation or collapse during printing or manufacturing. Different types of triangular facets have different shape characteristics. By classifying, more effectively extracting and utilizing these features to enhance the representational power of images or models. In image processing, the classification information of triangular patches is used to extract feature information such as edges, contours, or textures.



**Fig. 2.** Initial patch classification.

Different types of triangular facets have different shape characteristics. By classifying, more effectively extracting and utilizing these features to enhance the representational power of images or models. In image processing, the classification information of triangular patches is used to extract feature information such as edges, contours, or textures.

Since each facet that constitutes the initial aerial electric power tower image is a triangular facet, thus the facets are categorized into four classes in this paper. As shown in [Figure 2](#), the red rectangular dots indicate the set of feature points  $Q_w(x, y)$  directly access to vertices with valid texture information.

For individual vertices that cannot obtain effective texture due to occlusion, deformation, etc. [20], the approximate texture information of these points will be solved using a texture mapping method based on the one-ring neighborhood of vertices and facet classification. In this paper, the approximate texture information is obtained by solving the approximate texture coordinates of a vertex [21]. During the texture mapping process of aerial images of power towers, due to factors such as occlusion and deformation, the texture information of some feature points cannot be directly obtained. To solve this problem, a texture mapping method based on vertex ring neighborhood and patch classification is adopted. Vertex ring neighborhood is an important concept in graph theory, computer graphics, and computer vision, which describes the set of vertices directly adjacent to a given vertex in a graph or mesh structure, as well as the connection relationships between these vertices and the given vertex. In aerial images of power towers, the vertex ring neighborhood of feature points contains other directly connected feature points, and the spatial relationships and geometric characteristics between these feature points are crucial for extracting representative and robust texture features. Patch classification technology further enhances the accuracy of texture mapping. In aerial images of power towers, each patch is treated as a patch and classified based on its geometric shape and texture features. By classification, the geometric shape and texture features of power towers can be more accurately described which plays a key role in the subsequent texture mapping process. When a feature point cannot directly obtain texture information due to occlusion or deformation, the texture information of other feature points in its vertex ring neighborhood is used for approximation. The distance between the feature point and each feature point in its vertex ring neighborhood, and assigning different weights based on the distance size. Then, using these weights and the texture information of adjacent vertices with known

texture coordinates, the approximate texture coordinates of the feature point are solved through a weighted average method. Firstly, feature points are extracted from aerial images of power towers and classified based on their geometric shape and texture features. For each feature point, construct its vertex ring neighborhood, that is, find other feature points directly connected to it. For feature points that directly obtain texture information, record their texture coordinates; for feature points that cannot directly obtain texture information, approximate solutions are obtained using the texture information of other feature points in their vertex ring neighborhood. Map the calculated texture coordinates onto the corresponding feature points to form a complete texture mapping map. After the texture mapping is completed, methods such as the Babbitt coefficient are used to remove mismatched points to ensure the accuracy of feature point matching. The approximate texture coordinates are solved as follows:

Assuming the vertices, the  $Q$  is the set of feature points of the aerial electric power tower image  $Q_w(x, y)$  in which a valid texture-informative vertex with residual information, which is contained in a one-ring neighborhood, is contained  $m$  vertex, denoted as  $q_j(1 \leq j \leq m)$ , the distance between vertex  $Q$  with the vertices  $Q_j$  is  $C_q$ , the smaller the value  $C_q$ , vertex  $Q$  the closer the texture coordinate values are to the vertices  $Q_j$  texture coordinates, which means the closer the texture information of the two vertices is. However, if only the texture coordinates of the two vertices are used, the  $C_q$  takes the minimum value when the vertex, the  $Q_j$  texture coordinates, to approximate the vertices  $Q$  texture coordinates, which will produce a large error. For this reason, in this paper, we synthesize the vertices, the  $Q$  in a one-ring neighborhood that have all vertices reasonable texture coordinate values ( $[0, 0] \sim [1, 1]$ ) (*i.e.*, texture information is known) to compute the approximate texture coordinates of the vertices  $Q$ .

Set the approximate texture coordinates for the vertices  $Q(ux, uy, uz)$  in the feature point set to  $Q(hx, hy)$ , which is solved for using the following equation  $Q(hx, hy)$ , which gives the vertex  $Q$  of all the approximate texture coordinates, to get all the texture features  $Q'$  of the one-ring neighborhood of the vertices of the representative electric power tower image. Among them,  $\alpha_q$  denotes the impact factor, the smaller the  $C_q$ , the larger the value of impact factor  $\alpha_q$ .

$$\alpha_q = \frac{C_q}{Q_w(x, y)} \quad (6)$$

$$Q(hx, hy) = \sum_{j=1}^m \alpha_q(Q_{jh_x}, Q_{jh_y}). \quad (7)$$

Among them,  $Q_{jh_x}$  and  $Q_{jh_y}$  are the  $j$ th approximate texture pixel coordinates, the  $j \in m$ .

The innovation and advantage of this method is that the problem is not solved by processing the texture image, but by calculating the approximate texture coordinates of the vertices in the feature point set  $Q_w(x, y)$ , and then obtaining their approximate texture information, so as to extract the geometric representative texture characteristic points of the aerial electric power tower image.

### 2.3 Texture feature point matching based on the similarity principle

Considering that the methods based on vertex one-loop neighborhood and face-sheet classification may involve a large number of computations and complex data processing, in order to improve the speed of final feature point matching, the relative stability of triangles is utilized to introduce them into the feature matching algorithm. In this paper, two triangles are constructed in the real-time graph and the reference graph, respectively, by the hexadecimal group, and the principle of similar triangles is utilized to design a new feature matching method. Firstly, using the method described in Section 2.2, extract texture feature points from both the real-time image and the reference image. These feature points usually have significant texture information and high contrast, which is beneficial for the subsequent matching process. Select three texture feature points from the real-time image and reference image respectively, and construct two triangles. According to the principle of similar triangles, if two triangles are similar, the proportion between their corresponding sides should be equal, and the degree between their corresponding angles should be equal. In practical applications, due to the scaling and rotation changes in images, matching is performed using the characteristic that these changes do not affect the similarity of triangles. Determine whether two triangles are similar by calculating their side length ratio and angle relationship. If they are similar, then the points in the real-time graph match the points in the reference graph. By utilizing the relative stability of triangles for feature matching, complex calculations and data processing processes are avoided, thereby improving the Matching Speed (MS). The principle of similar triangles makes this method highly adaptable to image scaling and rotation changes, and to some extent overcomes the interference of image deformation and noise. By constructing multiple triangles for matching verification, the accuracy of matching can be further improved and the false matching rate can be reduced.

Combined with the principle of similar triangles, assuming in the real-time graphs  $\phi$ , the texture feature points of the three vertices obtained by the method mapping in Section 2.2 are  $Q_j \in Q''$ ,  $Q_k \in Q''$  and  $Q_v \in Q''$ , the matched points in reference chart  $P$  are, respectively  $P_i$ ,  $P_h$  and  $P_u$ . Then in the real-time graph  $\phi$ , regardless of the presence or absence of proportional and rotational changes, the triangle formed by  $Q_j$ ,  $Q_k$  and  $Q_v$ , with the triangle formed by  $P_i$ ,  $P_h$  and  $P_u$  should all be similar.

The sides of the two triangles can be expressed separately as:

$$\begin{cases} \Delta Q_{jk} = Q_j - Q_k \\ \Delta Q_{jv} = Q_j - Q_v \\ \Delta Q_{kv} = Q_k - Q_v \end{cases} \quad (8)$$

$$\begin{cases} \Delta P_{ih} = P_i - P_h \\ \Delta P_{iu} = P_i - P_u \\ \Delta P_{hu} = P_h - P_u \end{cases} \quad (9)$$

Then, by the principle of similar triangles, there should have:

$$\frac{\|\Delta Q_{jk}\|}{\|\Delta P_{ih}\|} = \frac{\|\Delta Q_{jv}\|}{\|\Delta P_{iu}\|} = \frac{\|\Delta Q_{kv}\|}{\|\Delta P_{hu}\|}. \quad (10)$$

For each hex  $(Q_j, P_i, Q_k, P_h, Q_v, P_u)$ , define the similarity of the two triangles, set  $\gamma_{jikh}(v, u)$  is when the aerial electric power tower image feature points, the  $Q_j$  and  $P_i$  paired and  $Q_k$  and  $P_h$  paired, the similarity of the triangle formed by  $Q_j, Q_k$  and  $Q_v$  with the triangle formed by  $P_i, P_h$  and  $P_u$ .

Hypothesis  $\gamma_{jikh}(v, u)$  is zero, then it is denoted the triangle formed by an aerial electric power tower image feature point  $Q_j, Q_k$  and  $Q_v$ , with the triangle formed by  $P_i, P_h$  and  $P_u$  are exactly similar, *i.e.*, they represent the pairs of feature points of the aerial electric power tower image  $(Q_v, P_u)$  should be given  $(Q_j, P_i, Q_k, P_h)$  with maximum support. Along with  $\gamma_{jikh}(v, u)$  increase, the degree of support should decrease.

Make the degree of support for  $(Q_v, P_u)$  to  $(Q_j, P_i, Q_k, P_h)$  is:

$$\eta(\gamma_{jikh}(Q_v, P_u)) = \frac{1}{1 + |\gamma_{jikh}(Q_v, P_u)|^2}. \quad (11)$$

In the formula,  $\gamma_{jikh}(Q_v, P_u)$  represents the similarity difference between the triangle composed of feature points  $i, j, k, h$  and the triangle composed of feature points. The value of support  $\eta(\gamma_{jikh}(Q_v, P_u))$  is between 0 and 1, with a larger value indicating that the two triangles are more similar and have higher support.

Requires that when the aerial electric power tower image feature points, the  $Q_j$  and  $P_i$  paired, and when  $Q_k$  and  $P_h$  paired,  $Q_v$  only matches the  $P_u$  with the greatest support for  $(Q_j, P_i, Q_k, P_h)$ , which can obtain the support degree  $\max \eta(\gamma_{jikh}(Q_v, P_u))$ . Take the average of the sums  $P_u$ , obtained the initial match metric value  $(Q_j, P_i, Q_k, P_h)$  is:

$$R^0(Q_j, P_i, Q_k, P_h) = \sum_{v \neq j, v \neq k} \max \eta(\gamma_{jikh}(Q_v, P_u)). \quad (12)$$

The initial matching metric value  $R^0(Q_j, P_i, Q_k, P_h)$  is calculated by averaging the support of all possible triangle combinations.

The initial matching measure of the feature point to the  $(Q_j, P_i)$  is:

$$R_*^0(Q_j, P_i) = \max R^0(Q_j, P_i, Q_k, P_h). \quad (13)$$

In the  $s$ th iteration, the  $(Q_v, P_u)$  to  $(Q_j, P_i, Q_k, P_h)$  the value of the matching metric, while relying on the difference in position between  $Q_v$  and  $P_u$ , and the initial match metric value  $R_*^{s-1}(Q_v, P_u)$  at the  $s - 1$ th iteration. In order to reflect the interaction of these two factors, the smallest of them is taken, so that the value of the matching measure is:

$$R_*^s(Q_v, P_u) = \max R^s(Q_j, P_i, Q_k, P_h). \quad (14)$$

The matching metric value  $R_*^s(Q_v, P_u)$  is the smaller of the positional difference and the initial matching metric value, which reflects the combined effect of positional and initial matching metrics.

In order to avoid the problem caused by the improper selection of the threshold  $c_s$ , the basic point relaxation algorithm can use the following method to determine whether the iteration of the matching measurement ends.

Definition:

$$c_s = \sum_{j=1}^n \sum_{i=1}^m \sum_{k \neq j}^n \sum_{h \neq i}^m |R^s(Q_j, P_i, Q_k, P_h) - R^{s-1}(Q_j, P_i, Q_k, P_h)|. \quad (15)$$

Among them,  $n$  represents the total number of feature points in the reference map. Then when  $c_s < \lambda$ , iteration terminated.  $\lambda$  is a predetermined very small positive number.

In order to avoid the problem of mis-matching, this paper introduces the Bartlett's coefficient to eliminate the mis-matched points. The Babbitt coefficient calculates the similarity between two feature points by comparing their kernel weighted Hu histograms. Hu histogram is a commonly used image feature representation method, which is calculated based on the moment features of the image and has rotation, scaling, and translation invariance. Kernel weighting is used to enhance the local information of feature points and improve the accuracy of matching. When the kernel weighted Hu histograms of two feature points are more similar, their Babbitt coefficients are closer to 1. This means that the more consistent the representation of these two feature points in the image, the more likely they are to be the correct matching pair. On the contrary, if the Babbitt coefficient of two feature points is low, their similarity is low and they are mismatched feature points. In order to eliminate mismatched feature points, a threshold is set to distinguish between feature points with high similarity (correct matching) and feature points with low similarity (false matching). Due to differences in feature point distribution and similarity between different images, using a fixed threshold is not always effective. To solve this problem, the dynamic threshold is calculated based on the distribution of the Babbitt coefficient. Firstly, calculate the Babbitt coefficients of all matching feature points and sort them. Then, observing the sorted Babbitt coefficient array, it was found that the Babbitt coefficients of mismatched feature points are usually much lower than those of other feature points. Therefore, utilizing this feature to construct a template and calculate the gradient of the Babbitt coefficient array. The gradient reflects the rate of change of the Babbitt coefficient array, that is, the change in feature point similarity. By selecting the maximum gradient as the threshold, it effectively eliminates the mis-matched points with Babbitt coefficients much lower than other feature points. This is because the maximum gradient usually marks a significant change in the similarity distribution, which is the boundary between correct and incorrect matches.

The value of the Babbitt coefficient lies within 0-1, and the higher its value, the higher the similarity of the feature points of the aerial electric power tower image, the calculation method of the Babbitt coefficient  $\sigma$  is:

$$\sigma = c_s \sum_o^O \sqrt{\phi_o P_o}. \quad (16)$$

Among them,  $\phi_o$ ,  $P_o$  represent the kernel weighted Hu histogram of the overlapping area of feature points after the initial matching of similar feature points of aerial electric power tower images.  $o \in O$  is the total number of matched aerial electric power tower image feature points. The value of the Babbitt coefficient  $\sigma$  is between 0 and 1, with higher values indicating that the two feature points are more similar.

After the calculation of  $\sigma$ , it is necessary to calculate the threshold of the comparison, if  $\sigma$  is less than the threshold value, it indicates that the feature points are matched incorrectly and need to be eliminated. The setting of the threshold is the key to the rejection of the mis-matched points. In this paper, the dynamic threshold is calculated according to the distribution of  $\sigma$ . In order to calculate the threshold value, the bar coefficients of all the matched feature points are sorted, and the Babbitt coefficients of the mismatched feature points are much lower than the Babbitt coefficients of other feature points. After sorting the similarity of the feature points, the gradient of the array of Babbitt coefficients reflects the correctness or incorrectness of the feature point matches. To compute the gradient, the template  $L$  is constructed:

$$L = [1, 1, 1, 1, 1, -1, -1, -1, -1, -1]. \quad (17)$$

Computing the gradient  $E$  of the array of Babbitt coefficients:

$$E(o) = \sum_o L\sigma. \quad (18)$$

Selecting the maximum gradient as the false match rejection threshold:

$$\tau = \max \{E(o)\}. \quad (19)$$

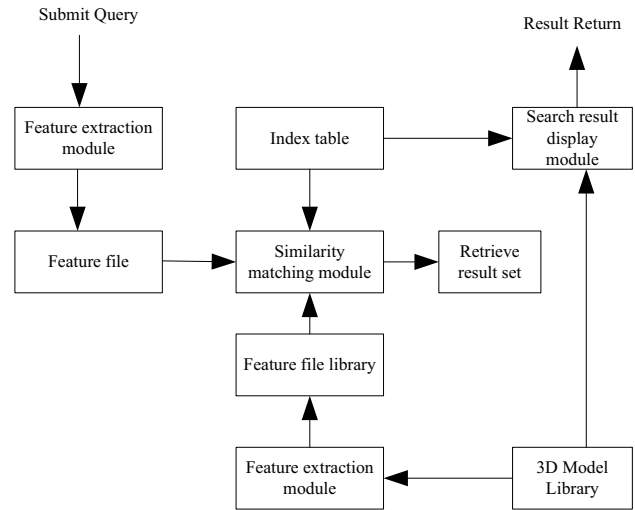
When false matches are rejected, if  $\sigma < \tau$ , retaining feature pairs,  $\sigma > \tau$ , culling feature point pairs. Gradient  $E$  reflects the rate of change of the Babbitt coefficient array, that is, the change in feature point similarity. By selecting the maximum gradient as the threshold, we eliminate the mis-matched points with Babbitt coefficients much lower than other feature points.

## 3 Experimental analysis

### 3.1 Experimental design

On the basis of 3D model library, a simple 3D model retrieval prototype system based on B/S structure is designed and implemented experimentally. The main structure of the prototype system is shown in Figure 3.

In Figure 3, after the query model submitted by the user is uploaded to the server, the shape features of the 3D model are firstly extracted by the feature extraction module and stored as a feature file, and then the similarity matching module utilizes the method in this paper to match the feature file with the feature file in the feature file library in the server for similarity calculation to get the retrieval result set. This retrieval result set is actually just an index



**Fig. 3.** A prototype system for 3D model retrieval based on B/S structure.

set of 3D models, so in the retrieval result display module, it needs to access the index table and the 3D model library to form a display page that can be browsed by the user, and finally the result page is returned to the user. The aerial power tower image dataset is derived from actual aerial projects and covers power tower images from different weather conditions, time periods, and geographical locations. The images of various types of power towers, such as straight line power towers and corner power towers, as well as background environments of different levels of complexity. The 3D model library contains 3D models corresponding to aerial images of power towers, which have been finely constructed to accurately reflect the structural and texture features of power towers. The 3D model is stored in a common 3D file format for subsequent feature extraction and matching operations. The feature point dataset obtained a large amount of feature point data by extracting feature points from aerial images of power towers. The feature point dataset contains information such as the position, direction, scale, and corresponding texture features of feature points. Used for training matching algorithms, optimizing parameter settings, and evaluating matching performance. In order to test the performance of the method under different lighting conditions, multiple aerial images of power towers with uneven lighting were randomly selected as the test set. These images contain areas that are too bright, too dark, and alternate between light and dark, posing challenges for feature extraction and matching. In the feature extraction stage, appropriate image filtering algorithms were used to improve image clarity, and reasonable feature point detection parameters were set. Then, the method proposed in this article is used for feature point matching, and the Root Mean Square Error (RMSE) and Matching Success Rate (MSR) of the matching results are calculated. In order to test the scale invariance of the method, aerial images of power towers were scaled at different scales. By changing the size of the image, power tower



**Table 1.** Experimental parameters.

Parameter	Value/Range
Gaussian pyramid layers	Floors 3–5
Laplacian pyramid layers	Same number of layers as Gaussian pyramid
Direction filter group direction number	8 directions
Feature point extraction window size	5 × 5 pixels
Sliding window size	5 × 5 pixels
Template size for calculating Babbitt coefficient threshold	5 × 5
Selection of dynamic threshold for Babbitt coefficient	Select the value at the maximum gradient as the threshold
Feature file storage format	Binary, Text
Display method of search results	Image list

images captured at different distances were simulated. In the feature extraction stage, keep the feature point detection parameters unchanged to test the scale invariance of the method. Then, perform feature point matching and calculate the RMSE and MSR of the matching results. To test the rotational invariance of the method, aerial images of power towers were rotated at different angles. By rotating the image, simulated images of power towers taken from different perspectives. Maintain the feature point detection parameters unchanged during the feature extraction stage to test the rotational invariance of the method. Then, perform feature point matching and calculate the RMSE, MSR, and MS of the matching results. The experiment used high-performance computers for data processing and calculation, ensuring the smooth progress of the experiment. The experiment was developed based on open-source libraries such as Python and OpenCV, and implemented features extraction, matching, and performance evaluation. The B/S structure prototype system in the experiment was deployed within the local area network, ensuring the stability and security of data transmission. The specific experimental parameters involved are shown in Table 1.

To verify the effectiveness of the method proposed in this paper, a feature point matching experiment was conducted using the linear power tower shown in Figure 4 as an example. The aerial image is affected by weather and has poor imaging quality, which poses certain challenges.

As shown in Figure 4, the aerial image of this linear electric power tower is affected by the weather, resulting in poor imaging quality of the aerial image.

In the experiment, RMSE is used as the index to measure the matching accuracy. RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{\sum_{j=0}^m (x_j + y_j - x'_j)^2 + (x_j + y_j - y'_j)^2}{m}}. \quad (20)$$

Among them,  $(x_j, y_j)$  and  $(x'_j, y'_j)$  represent the coordinates of the texture feature points in the image to be matched, and the coordinates of the texture feature points with the same name in the reference image, respectively.  $m$  represents the number of feature points, the  $j \in m$ .

**Fig. 4.** Aerial images of straight electric power towers.

## 3.2 The methodology in this paper uses performance tests

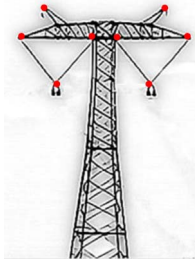
### 3.2.1 Similar feature point matching effect in uneven illumination

Uneven lighting is a common challenge in the process of retrieving 3D model maps from aerial electric power tower images. Uneven illumination may cause some parts of the image to be too bright or too dark, thus affecting the accuracy of feature extraction and matching. In order to test the performance of the proposed method under uneven lighting conditions, when extracting features from aerial images of power towers under uneven lighting conditions, multiple aerial images of power towers with uneven lighting were selected as the test set to ensure that the images contained the complete structure and texture information of the power towers. In the feature extraction stage, appropriate image filtering algorithms are used to improve image clarity, and reasonable feature point detection parameters are set. The image filtering effect is shown in Figure 5, and the labeling diagram of the extracted core feature points is shown in Figure 6.

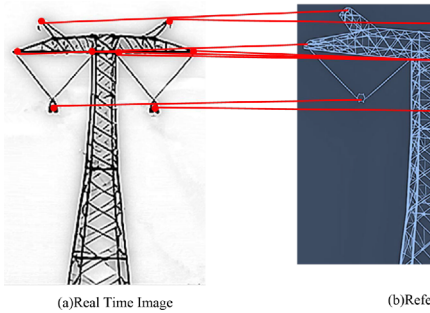
As shown in Figures 5 and 6, when the method in this paper is used to extract features from aerial electric power tower images, filtering can effectively improve the image clarity, and feature extraction of prominent positions such as corners and endpoints of electric power towers is accurate. Figure 7 shows the matching effect of similar feature points when the light is uneven.



**Fig. 5.** Aerial electric power tower image filtering effect diagram.



**Fig. 6.** Schematic diagram of marking after extracting the core feature points of the image.



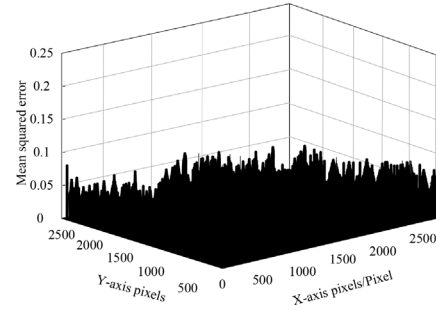
**Fig. 7.** Matching effect of similar feature points under uneven lighting conditions.

As shown in Figure 7, the method of this paper can accurately match the corresponding points in the 3D model map when retrieving the 3D model map of the aerial electric power tower image with uneven illumination. Perform similarity calculation between the extracted feature points and the feature file library in the server to obtain matching results. In order to quantify the matching accuracy, RMSE is used as the measurement index, and the calculation formula is shown in (20). Figure 8 shows the change of the mean square error of the matching result of the feature point location after matching similar texture feature points by this method in this case.

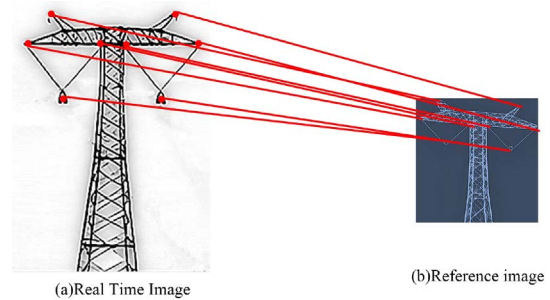
As shown in Figure 8, when the light is uneven, the method in this paper matches the similar feature points of aerial electric power tower images, and the mean square error of the matching result of the feature point position is within 0.1 Pixel, and the matching result is accurate.

### 3.2.2 Effectiveness of similar feature point matching with scale invariance

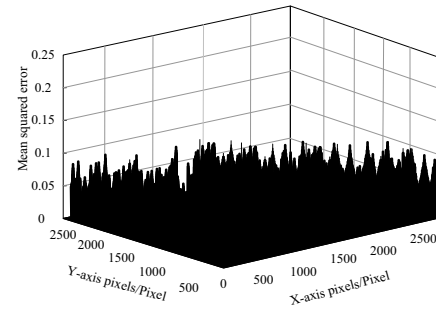
In order to test the scale invariance of the method proposed in this paper, aerial images of power towers were scaled at



**Fig. 8.** Mean squared error of similar feature point matching results under uneven illumination.



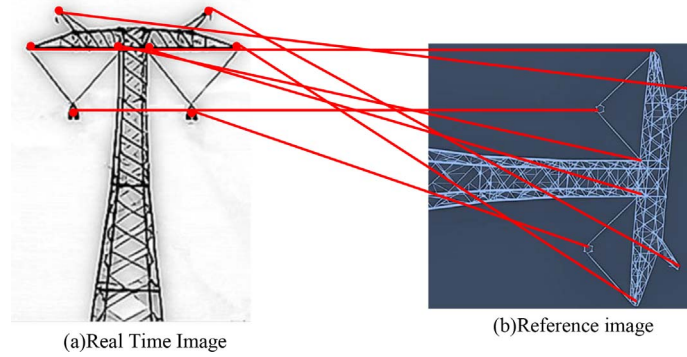
**Fig. 9.** Scale invariance experimental results.



**Fig. 10.** Mean squared error of similar feature point matching results with scale invariance.

different scales and similar feature point matching experiments were conducted. The experimental setup is similar to that of uneven lighting, but with the addition of image scale processing. In the feature extraction stage, keep the feature point detection parameters unchanged to test the scale invariance of the method. When the scale of 3D model of aerial electric power tower image occurs, the experimental results of similar texture feature point matching are shown in Figure 9, and Figure 10 shows the change of mean-square error of similar feature point matching results of this paper's method for scale invariance.

As shown in Figures 9 and 10, when the scale of the three-dimensional model of the aerial electric power tower image occurs, the matching result is accurate from the visual perspective after the scale invariant similar texture feature points are matched using the method in this paper;



**Fig. 11.** Experimental results of matching similar texture feature points.

From the perspective of mean square error test results, the matching error is still controlled within 0.1 Pixel, which can achieve high-precision matching of similar texture feature points in aerial electric power tower images.

### 3.2.3 The effect of matching similar feature points with rotational invariance

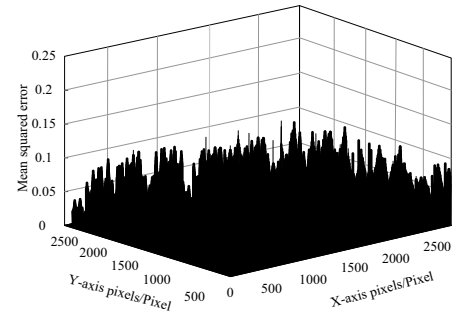
In order to test the rotational invariance of the method proposed in this article, aerial images of power towers were rotated at different angles and similar feature point matching experiments were conducted. In the feature extraction stage, the feature point detection parameters are also kept unchanged to test the rotational invariance of the method. Rotational invariance plays an important role in similar feature point matching, which can help the method to better deal with the image rotation problem and improve the accuracy and reliability of the matching. The experimental results of similar texture feature point matching of this paper's method in this case are shown in Figure 11, and Figure 12 is the change of the mean square error of the matching results of the similar feature points.

As shown in Figures 11 and 12, the rotation invariant similar feature point matching experiment, the rotation invariant similar texture feature point matching effect is ideal, the visual matching effect is accurate, and the mean square error of the matching result does not exceed 0.1 Pixel.

### 3.3 Similar feature point matching accuracy performance test

In order to deeply test the method of this paper in the same kind of matching method, whether there is the use of advantages, experiments selected aerial photography electric power tower image contour of a feature segment, which has two texture feature points (referred to as contour 1, contour 2), and at the same time selected two sampling contour of the same feature segment of the eight points, the curvature value of each point is shown in Table 2.

Set the curvature tolerance value within  $\pm 0.1$  to compare the matching accuracy of different methods. In feature point matching, the selection of curvature tolerance value needs to balance the accuracy and robustness of the matching. A too small tolerance value leads to overly strict matching, resulting in missing some correct matching pairs;



**Fig. 12.** Mean squared error of similar feature point matching results with rotational invariance.

**Table 2.** Details of contour curvature in aerial electric power tower images.

Target profile	Outline 1	Outline 2
0.824	0.813	0.837
0.769	0.764	0.649
0.802	0.798	0.698
0.688	0.675	0.544
0.453	0.469	0.237
0.362	0.298	0.447
0.639	0.628	0.369
0.537	0.548	0.786

Excessive tolerance values introduce incorrect matching pairs. The curvature tolerance value of  $\pm 0.1$  was found in experiments to improve the robustness of matching while ensuring a certain level of accuracy. In addition, through feature point matching experiments on a large number of images, it was found that  $\pm 0.1$  is a suitable threshold that provides stable matching results under different lighting conditions and image deformations, and provides a reasonable benchmark for comparison after comparing multiple tolerance values. The methods described in this article and literatures [5–8] were used to match similar feature points in the image under uneven lighting conditions. The

**Table 3.** Changes in RMSE of curvature matching for similar feature points in images.

Target profile	Proposed method		Reference [5] method		Reference [6] method		Reference [7] method		Reference [8] method	
	Outline 1	Outline 2	Outline 1	Outline 2	Outline 1	Outline 2	Outline 1	Outline 2	Outline 1	Outline 2
0.824	0.01	0.01	0.14	0.18	0.11	0.12	0.22	0.24	0.28	0.29
0.769	0.01	0.02	0.15	0.19	0.12	0.12	0.24	0.27	0.29	0.31
0.802	0.02	0.01	0.17	0.17	0.13	0.12	0.27	0.26	0.28	0.32
0.688	0.01	0.01	0.18	0.19	0.12	0.11	0.26	0.24	0.28	0.33
0.453	0.02	0.02	0.14	0.17	0.13	0.13	0.25	0.27	0.29	0.31
0.362	0.01	0.01	0.15	0.17	0.12	0.11	0.25	0.26	0.28	0.32
0.639	0.01	0.02	0.17	0.18	0.13	0.13	0.25	0.29	0.28	0.32
0.537	0.01	0.02	0.15	0.19	0.11	0.11	0.25	0.22	0.28	0.32

**Table 4.** Performance comparison under different conditions.

Different conditions	Method	RMSE	MSR/%	MS/images/s
Light variation	Proposed method	3.2	92.5	25
	Reference [5] method	4.5	85.0	20
	Reference [6] method	4.0	88.0	22
	Reference [7] method	3.8	90.0	23
	Reference [8] method	3.5	91.0	24
Scale scaling	Proposed method	2.8	94.0	24
	Reference [5] method	4.2	83.0	19
	Reference [6] method	3.9	86.0	21
	Reference [7] method	3.6	89.0	22
	Reference [8] method	3.3	92.0	23
Angle rotation	Proposed method	3.0	93.5	24
	Reference [5] method	4.8	80.0	18
	Reference [6] method	4.3	84.0	20
	Reference [7] method	3.9	87.0	21
	Reference [8] method	3.4	91.0	23

RMSE of the curvature of the feature points is shown in Table 3.

Analyzing the data in Table 3, it can be seen that, in the working condition of uneven illumination, the RMSE of feature point curvature of this paper's method, the literature [5] method, the literature [6] method, the literature [7] method, and the literature [8] method after matching the feature points of the similar texture of the image is significantly different, but after the use of this paper's method, the RMSE of the feature point curvature matching is in the range from 0.01 to 0.02, and the literature [5] method, Literature [6], literature [7], and literature [8], the RMSE of feature point curvature matching is larger than that of this paper's method, which indicates that this paper's method has a higher accuracy of feature point curvature matching under the condition of uneven illumination. In contrast, the RMSE of feature point curvature matching of the methods described in [5–8] are larger than

that of this method under the same conditions, which indicates that the matching accuracy of these methods is relatively low. Uneven lighting is a common problem in image processing, which can cause certain areas in the image to be too bright or too dark, thereby affecting the accurate extraction of feature points. In this case, the position, size, and shape of the feature points may change, which in turn affects the subsequent matching process. This article adopts a certain lighting preprocessing technique or a robust feature extraction algorithm to stably extract feature points under uneven lighting conditions, while maintaining the accuracy of their curvature and other attributes. This article adopts advanced matching algorithms to maintain high matching accuracy under uneven lighting conditions. In addition, this method also incorporates other information to improve the robustness of matching. This method has undergone extensive experimentation and optimization, maintaining high performance in various



scenarios and conditions. In addition, this method also considers other factors that may affect the matching accuracy, thereby improving its generalization ability. In practical applications, this method maintains high matching accuracy while also having high computational efficiency and low memory usage. This gives the method proposed in this article greater advantages in practical applications.

In order to further demonstrate the superiority of our method and verify its comprehensive performance under different lighting conditions, scales, and rotation angles. Select multiple sets of aerial images of power towers and perform light changes, scale scaling, and angle rotation processing separately. Keep the parameter settings of each method unchanged to fairly compare their performance. Similarly, the methods from references [5–8] were selected as comparison objects. In addition to RMSE, MSR, and MS are introduced as evaluation metrics. The experimental results are shown in Table 4.

According to the analysis in Table 4, the method proposed in this paper has significant performance advantages in processing aerial images of power towers, especially in terms of comprehensive performance under different lighting conditions, scaling at different scales, and rotation angles. Among them, under the condition of lighting changes, the RMSE of our method is the lowest, the MSR is the highest, and the MS is relatively fast, indicating that its performance under lighting changes is better than other methods. Under scaling conditions, our method also performs well, with the lowest RMSE, highest MSR, and stable MS. Under the condition of angular rotation, the method proposed in this paper still maintains a high performance advantage, with RMSE and MSR outperforming other methods, and the MS remains unchanged. This is mainly due to the advanced algorithms and optimized parameter settings used in this method, which have a much higher matching success rate than other methods, indicating its higher stability and accuracy in processing complex images. Although there are differences in MS between different methods, the method proposed in this paper maintains a relatively fast MS while ensuring high performance, which is of great significance for real-time requirements in practical applications. Therefore, the method proposed in this article has broad application prospects in practical applications.

## 4 Conclusion

The method of matching similar feature points in aerial electric power tower images based on a vertex ring neighborhood is an efficient feature extraction and matching method designed specifically for aerial electric power tower images. This method deeply explores the local features and structural information of electric power tower images, extracts geometrically representative electric power tower feature points through a unique vertex one ring neighborhood strategy, and uses accurate similarity measurement to achieve effective feature point matching.

In summary, this method demonstrated the following significant conclusions in the experiment:

1. By utilizing a vertex ring neighborhood, this method can accurately capture local features and structural details of electric power tower images, ensuring that the extracted feature points are both representative and stable.
2. When facing the problem of matching similar texture feature points in complex situations such as uneven lighting, scale changes, and rotation changes, this method performs well and can strictly control the position matching error of texture feature points within 0.1 Pixel.
3. Compared with similar methods, under the same conditions, the method proposed in this paper shows a lower RMSE in feature point curvature matching of aerial electric power tower images, proving its higher applicability and reliability in processing complex scenes such as aerial electric power tower images.

However, although this method has achieved significant results in experiments, there are still some potential limitations. Although this method performs well in feature extraction and matching, its computational complexity is relatively high. This can result in longer processing times when dealing with large-scale image datasets. Some parameters in this method may have a significant impact on the results. Therefore, in practical applications, it is necessary to carefully adjust these parameters to achieve the best results. To address the issue of high computational complexity, attempt to optimize the algorithm structure and reduce unnecessary computational overhead. To reduce parameter sensitivity, a self-adaptive parameter adjustment mechanism is studied. This mechanism automatically adjusts parameter values based on the characteristics of the input image, thus achieving the best matching effect without manual intervention. In addition to vertex neighborhood features, other types of features such as color features and texture features can also be fused to improve the accuracy and robustness of matching. By fusing multiple features, useful information in the image can be further explored to improve matching performance.

### Funding

The authors received no financial support for the research.

### Conflicts of interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and publication of this article.

### Data availability statement

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

### References

- 1 Forkan A.R.M., Kang Y.B., Jayaraman P.P., Liao K., Kaul R., Morgan G., Ranjan R., Sinha S. (2022) CorrDetector: a framework for structural corrosion detection from drone images using ensemble deep learning, *Expert Syst. Appl.* **193**, 116461.1–116461.12.

- 2 Ahmed M.D., Faiyaz M.J.C., Sanyal A. (2022) Inspection and identification of transmission line insulator breakdown based on deep learning using aerial images, *Electr. Power Syst. Res.* **211**, 1–15.
- 3 Ery A., Lin Z. (2022) Template matching and change point detection by M-estimation, *IEEE Trans. Inf. Theory* **68**, 1, 423–447.
- 4 Ahmad A. (2021) Feature aggregation with attention for aerial image segmentation, *IEEE Sens. J.* **21**, 23, 26978–26984.
- 5 Sourabh P., Udaysankar D., Yashwanth N., Yogeswara R. (2022) An efficient SIFT-based matching algorithm for optical remote sensing images, *Remote Sens. Lett.* **13**, 10/12, 1069–1079.
- 6 De L.R., Cabrera-Ponce A.A., Martinez-Carranza J. (2021) Parallel hashing-based matching for real-time aerial image mosaicing. *J. Real-Time Image Process.* **18**, 1, 143–156.
- 7 Meddeber L., Zouagui T., Berrached N. (2021) Efficient photometric and geometric stitching approach for remote sensing images based on wavelet transform and local invariant, *J. Appl. Remote Sens.* **15**, 3, 034502-1–034502-31.
- 8 Ekanayake E.M.C.L., Lei Y. (2021) Crowd estimation using key-point matching with support vector regression, *IET Image Process.* **15**, 14, 3551–3558.
- 9 Luo L. (2024) Research on inspection path planning of power tower, *Manuf. Autom.* **46**, 1, 56–62.
- 10 Harbi J.S., Najm Z.N. (2022) Image decomposition using deep variation priors system, *Nonlinear Opt. Quantum Opt.* **56**, 3/4, 193–204.
- 11 Sumon K.B., Deepak S., Arindam B. (2022) A 51.3-TOPS/W, 134.4-GOPS in-memory binary image filtering in 65-nm CMOS, *IEEE J. Solid-State Circuits* **57**, 1, 323–335.
- 12 Kotov V.M. (2021) Two-dimensional image edge enhancement using a spatial frequency filter of two-color radiation, *Quantum Electron.* **51**, 4, 348–352.
- 13 Devulapalli S., Krishnan R. (2021) Remote sensing image retrieval by integrating automated deep feature extraction and handcrafted features using curvelet transform, *J. Appl. Remote Sens.* **15**, 1, 016504-1–016504-8.
- 14 Li W., Resmerita E., Vese L.A. (2021) Multiscale hierarchical image decomposition and refinements: qualitative and quantitative results, *SIAM J. Imaging Sci.* **14**, 2, 844–877.
- 15 Tarsitano F., Bruderer C., Schawinski K., Hartley W.G. (2022) Image feature extraction and galaxy classification: a novel and efficient approach with automated machine learning, *Mon. Not. R. Astron. Soc.* **511**, 3, 3330–3338.
- 16 Geet L., Chitta R., Jialei C., Hao Y., Chuck Z. (2023) Convolutional neural network-assisted adaptive sampling for sparse feature detection in image and video data, *IEEE Intell. Syst.* **38**, 1, 45–57.
- 17 Dong F.J., Yuan Y. (2023) Simulation of local feature spot detection of panoramic visual images in complex scenes, *Comput. Simul.* **40**, 7, 168–171 189.
- 18 Salmi A., Hammouche K., Macaire L. (2021) Constrained feature selection for semisupervised color-texture image segmentation using spectral clustering, *J. Electron. Imaging* **30**, 1, 13014.1–13014.28.
- 19 Kesav O., Homa R.G.K. (2021) Automated detection system for texture feature based classification on different image datasets using S-transform, *Int. J. Speech Technol.* **24**, 2, 251–258.
- 20 Nouri M., Baleghi Y. (2023) An active contour model reinforced by convolutional neural network and texture description, *Neurocomputing* **528**, 125–135.
- 21 Kesav O., Homa R.G.K. (2021) Automated detection system for texture feature based classification on different image datasets using S-transform, *Int. J. Speech Technol.* **24**, 2, 251–258.