IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 26, 2019, accepted October 8, 2019, date of publication October 17, 2019, date of current version October 31, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2948088

A Local Feature Descriptor Based on SIFT for 3D Pollen Image Recognition

ZHUO WANG[1], WENZHENG BAO[1], DA LIN[1], AND ZIXUAN WANG[2]
[1]School of Information and Electrical Engineering, Xuzhou University of Technology, Xuzhou 221018, China
[2]School of International Studies, Sun Yat-sen University, Guangzhou 510275, China

Corresponding author: Zhuo Wang (wangzhuo@xzit.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 41675155 and Grant 61902337, in part by the Qinglan Project of Jiangsu Province in 2018, in part by the Natural Science Research of the Jiangsu Higher Education Institutions of China under Grant 18kja520012, in part by the Xuzhou Science and Technology Plan Project under Grant kc19197, and in part by the Cultivation Project of Xuzhou University of Technology under Grant xky2018126.

ABSTRACT Biological particle automatic classification is an important issue in index tasking for people with pollen hypersensitivity. This paper attempts to present a local feature extraction method based on SIFT for automatic 3D pollen image recognition. In order to solve major issues in previous studies, high rate of redundant information, high feature dimensions and low recognition rate should be taken into account. Therefore, this work focuses on a four-part novel approach, including constructing 3D Gaussian pyramid to obtain muti-scale pollen images, computing the local differential vector to explore local key points, filtering the key points by inter-layer contrast, and extracting the statistical histogram descriptor of the key points as discriminant feature for automatic classification of 3D pollen images. Experiments are performed on three standard pollen image datasets including Confocal, Pollenmonitor and CHMontior. It is concluded that the descriptor can effectively describe the pollen image and is robust to the rotation, translation and scaling of the image.

INDEX TERMS Scale invariant feature transform, local feature, pollen recognition, 3D image.

I. INTRODUCTION

With the improvement of the global environment in recent years, coverage of plants which easily release highly allergenic pollen, such as poplar, elm, oriental plane, etc. has increased rapidly [1]–[3]. Therefore, the high concentration of allergenic pollen in the air may undermine the health of those who are pollen hypersensitivity [4]. Thus researchers have been studying to alleviate the potential threat. According to some former researches about real-time monitoring of pollen gains' type, concentration and propagation path can effectively help us to forecast the highly allergenic pollen concentration [5]–[8]. The principal tasks of this method consist of classifying and counting different kinds of pollen gains which require highly skilled experts to obtain accurate identification results. This work, requiring much labor input, is arduous, time consuming and lack of consistent quality control. Some researchers [9]–[14] believe that it is an effective improvement to use computer vision and pattern recognition technique for automated pollen image classification since intrinsic digitalized features, such as shape, texture, geometric and statistical features, can be easily captured when they are obtained by microscopes [15]–[19].

Methods on automatic recognition and classification of pollen grains have been developing since Flenley initially proposed the need for automated palynology in 1996 [20]. Achievements in feature extraction for automatic pollen recognition can be reviewed with two aspects: planar features in 2D images and stereoscopic features in 3D images. Prior researches generally confirm that planar features can effectively describe the shape and structures of two-dimensional pollen images and usually have good rotation invariance [21]–[24]. Rodriguez-Damian *et al.* achieved 86% recognition rate in the dataset of the algorithm based on the combination of shape and texture analysis, which has been widely utilized in Urticaceae family. But the method was only applied to a single genus [25]. Lozano-Vega *et al.* considered about aperture features, and analyzed the effectiveness of shape features, texture features and aperture features (STAF). Meanwhile, they designed and developed multithreading functionality to decrease the processing time in large pollen dataset [26]. Marcos *et al.* evaluated different kinds of texture

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

extraction method which were widely used in pollen recognition, and combined GLCM (gray-level co-occurrence matrix) and DTM (discrete Tchebichef moments) to identify pollen images, and achieved a better recognition rate compared to single texture at the cost of increased time complexity [27]. According to empirical evidence, the methods above have their own advantages, but there are also many drawbacks, especially the considerable redundant information and the limitations of ignoring inner structure of pollen grains, which would easily lead to the decrease of identification efficiency. As for the stereoscopic features, some research substantiates the belief that these types of features can effectively describe the statistical characteristics of pollen images and have good robustness to rotation transformation, scale transformation and affine transformation. In addition, some categories of pollen grains have similar surface structure which would easily lead to misclassification in 2D cases. However, these types of pollen grains can be recognized in 3D cases owing to the significant differences in inner structure. Figure 1 shows two categories of pollen grains which have similar surface but different inner structures.

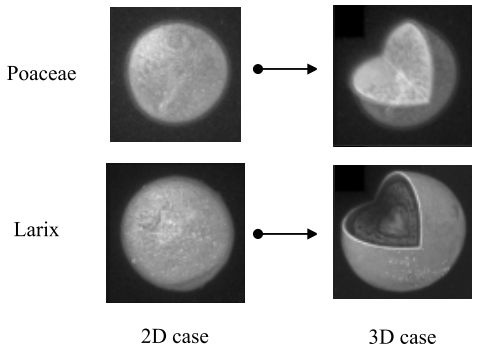

FIGURE 1. Examples of two typical categories of pollen grains in different dimension.

Ronneberger *et al.* extracted 14 invariant gray-scale features based on an integration over the Euclidian 3D transformation group with nonlinear kernels, using support vector machine to classify the 26 most important German pollen taxa, and achieved 82% recognition rate in Confocal pollen dataset [28]; Kim *et al.* presented a 3D refractive-index imaging and a quantitative analysis method of pinus pollen grains, and achieved quantitative 3D information of pollen grains using optical diffraction tomography while leaving the problem of time complexity unsolved [29]; Xie *et al.* developed a feature extraction method of Fourier descriptor based on the SC-Zernike moments(Zernike moments in spherical coordinates) for 3D pollen images recognition. The simulation results indicated that the descriptor was robust to the rotation, translation and scaling of 3D pollen images and achieved 90% recognition rate on Confocal dataset. But there is still room for improvement in the classification of deformed or clumped pollen images [30]. In summary, these descriptors can partly describe the 3D pollen image. However, redundant information caused by high dimension of 3D data is still a mainly contributory factor of low computational efficiency. Thus a highly efficient 3D descriptor extracting method should be proposed [31]. With recent development of machine learning algorithms in image processing and recognition [32]–[35], some achievements have been accomplished in automatic pollen recognition domain. Daood *et al.* combined stacks of multifocal images, convolutional and recurrent neural networks to learn the optimal features and recognized a pollen grain as a sequence of multifocal images acquired by an optical microscope [36]. Khanzhina *et al.* proposed a deep learning solution of 11 species pollen grain images based on a convolutional neural network(CNN) for classification, feature extraction and image segmentation [37]. Sledevi *et al.* found out a sufficient configuration of CNN required for implementation on low-cost FPGA to classify pollen bearing bee images [38]. Although the machine learning algorithms have achieved satisfactory classification results in laboratory pollen databases, there are still some non-ignorable drawbacks. High-definition image is required in convolutional neural network due to the processing procedure of pooling layer. However, the majority of auto-collected pollen images are excessively blurred for pooling. Thus recognition results may be unsatisfactory in such cases. Besides, copious training images are necessary for CNN to ensure its classification capacity, which makes it inapplicable for some small pollen datasets.

Compared with traditional images, pollen images have the property of distinct texture and conspicuous image corner in different scales. According to these qualities, we take both planar features and structural information into consideration, and present the local key point descriptor based on scale invariant feature transform(SIFT) for 3D pollen recognition called spatial local key point(SLKP). Instead of applying difference of Gaussian in SIFT, our method uses the local differential vector to explore local key points. Besides, another supporting vector is introduced to determine the direction of z-axis in the procedure of rotation transform of traditional SIFT. Experimental results on the standard pollen image dataset show that this method can effectively improve the recognition rate and has low complexity. Figure 2 shows the flowchart of the proposed method.

II. METHODS

A. CONSTRUCTION OF 3D GAUSSIAN PYRAMID

In this paper, we introduce a Z-axis vector to assist in detecting key points of 3D pollen image. The first step of key point detection is to identify locations and angle points that can be repeatably assigned under different scales of the same 3D pollen target. In order to obtain three-dimensional pollen image in different scales, 3D Gaussian scale space based on traditional Gaussian scale space should be constructed.

It has been shown by Koenderink that under a variety of reasonable assumptions the only possible scale-space kernel is the Gaussian function [39]. Therefore, the scale space of 3D pollen image is defined as a function, $L(x, y, z, \sigma)$, that is

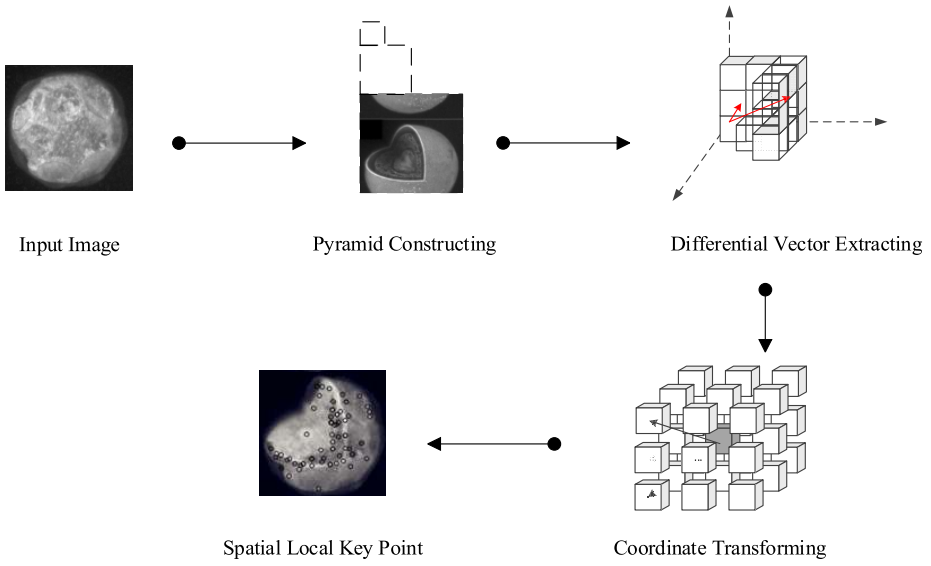

FIGURE 2. Flowchart of the proposed method.

produced from the convolution of a variable-scale Gaussian, G(x, y, z, σ), with an input image, I(x, y, z):

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

where, * is the convolution operation in x, y and z, and

$$G(x, y, z, \sigma) = \frac{1}{(2\pi)^{3/2}\sigma^3} e^{-(x^2+y^2+z^2)/2\sigma^2} \tag{2}$$

In traditional scale-invariant feature transform, difference of Gaussian function was used to detect the local extrema as the key-point. But in 3D cases, the rapid increase of the algorithm complexity and the high density data may lead to computational inefficiency. To efficiently detect stable key-point location in scale space of 3D pollen images, we have proposed using local differential vector instead of computing difference of Gaussian in SIFT.

B. EXTRACTION OF LOCAL DIFFERENTIAL VECTOR

The main motive of local differential vector is to detect the local maximum points of gradient. The extraction procedure is as follows:

Step 1: Dividing the image into blocks in each layer of the 3D Gaussian image pyramid. The size of the block is related to the scale factor k, which indicates the range of each block. Computing the mean value of pixels in each block as the sampling value.

Step 2: Dividing the sampling values into sampling center and sampling points as the output. Figure 3 indicates the sampling model in 3D neighborhood.

Step 3: Setting a Cartesian coordinate system based on each block. Calculating the differential vector between sampling center and sampling points in such coordinate system. The differential vectors are divided into positive group and

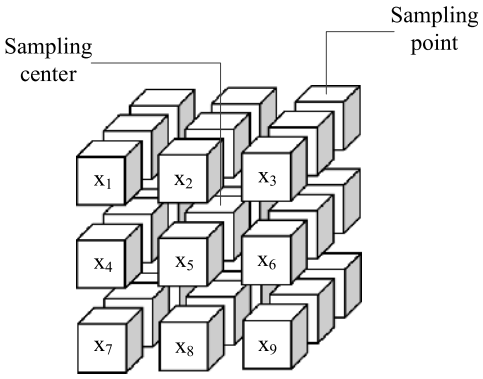

FIGURE 3. Sampling model in 3D neighborhood.

negative group according to the difference of gray value:

$$V^+ = \sum_i^G (v_c - v_i) D(v_c, v_i)$$
$$V^- = \sum_i^G (v_c - v_i) D(v_i, v_c) \tag{3}$$

where, G represents the related neighborhood of current sampling center. D(v_i,v_c) can be computed from the difference between sampling center v_c and sampling point v_i:

$$D(v_1, v_2) = \left\{ \begin{array}{l} 1, v1 \geq v2 \\ 0, else \end{array} \right\} \tag{4}$$

The norm of each differential vector depends on the difference between the sampling value and the center value. Direction of differential vector comes from the direction of the unit vector of sampling points under such coordinate system. Therefore, positive differential vector V^+ and negative differential vector V^- are extracted as the output of

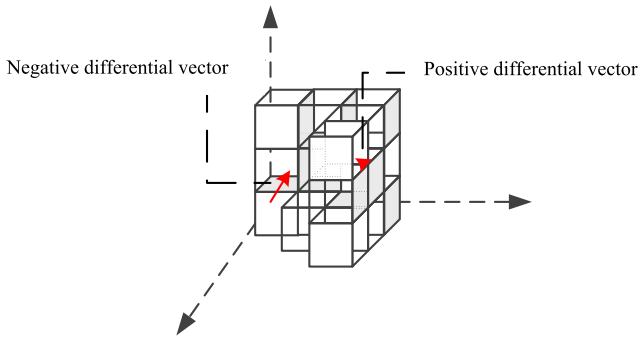

FIGURE 4. Extraction of positive & negative differential vector.

this step. Figure 4 shows the positive differential and the negative differential vector of the block.

C. DETECTION OF LOCAL KEY POINT

Just as we determined the differential vectors of each block, so we must determine the trend of the gray level in the image domain by positive differential vector and negative differential vector. In order to detect the stable local key point, gradient vector of each block in different layers is computed as follows:

$$V_{loc,\sigma} = V^{-}_{loc,\sigma} - V^{+}_{loc,\sigma} \tag{5}$$

where, loc and σ indicate the location and layer of the current block, and then we can obtain the norm of gradient vector as:

$$|V_{loc,\sigma}| = \sqrt{V_x^2 + V_y^2 + V_z^2} \tag{6}$$

Once the gradient vector has been obtained by calculating the difference of pixels to its neighbors as the result of this section, the next step is to detect the precise location of local key point by following procedures by inputting gradient vectors.

The first step is to detect the region of interest (ROI). The blocks which have greater norm of gradient vector than specific threshold are chosen:

$$Inst_\sigma = \{f_{(loc1,\sigma)}, f_{(loc2,\sigma)}, \ldots, f_{(loc3,\sigma)} || V_{loc,\sigma}| > Thr\} \tag{7}$$

where, $f_{(loc,\sigma)}$ represents the block at the corresponding location in the blocked image, and Thr is the threshold of the norm. Such method is used for the detection of ROI in the remaining layers of the Gaussian pyramid:

$$V_{n\sigma} = V^{+}_{loc,n\sigma} - V^{-}_{loc,n\sigma} | f_{loc} \in Inst \tag{8}$$

The final step is to screen the results by comparing gradient vectors of ROI, and marking the location where all the blocks in every layer have stable local structure. Then we select the regions from the marked area which have similar gradient vectors at the same location in different layers as the local key points (Figure 5).

$$I = \{f_{loc} | V_\sigma \approx V_{2\sigma} \approx \ldots \approx V_{n\sigma} | f_{loc} \in Inst\} \tag{9}$$

The blocks contained in the set I are the outputs of local key points mentioned above in this section.

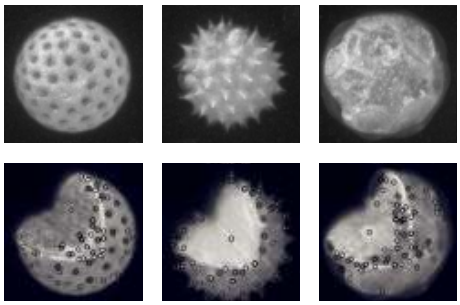

FIGURE 5. Examples of local key points obtained by proposed method.

D. DESCRIPTOR REPRESENTATION

After the images are described with local differential vectors, the next step is to solve the rotation problem of the method above. Because of the particularity of pollen images, requirement for robustness of rotation in such extraction method should be emphasized. In three-dimensional cases, rotation problem of pollen gains becomes far more complicated than traditional images. Therefore, we introduce a method, namely rotation invariant feature transform, into the extraction procedure of gradient vectors mentioned above in expectation of more stable features.

In three-dimensional cases, translation transformation and scale transformation can be processed easily as extension of 2D Cartesian system, but rotation transformation is relatively more complicated. In this paper, 3D rotation matrix is used to change the vectors from one system to another as follows:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} cos\alpha_x & cos\beta_x & cos\gamma_x \\ cos\alpha_y & cos\beta_y & cos\gamma_y \\ cos\alpha_z & cos\beta_z & cos\gamma_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = Trs \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{10}$$

where (X, Y, Z) is the local differential vector under previous coordinate and (X', Y', Z') is the rotated vector. α, β, γ are used to describe the angles between the previous axes and the rotated axes. According to Lowe's theory, the normalization of coordinate system should be taken into account to enhance the robustness of rotation transform. In traditional SIFT, a consistent orientation is assigned to each keypoint based on local image properties. The descriptor can be presented based on this orientation and therefore achieve invariance to image rotation. Similarly, two consistent orientations need to be assigned under the 3D coordinate system. Therefore, an assistant orientation should be set to determine the second axis' direction. As it is shown in figure 6, twenty-six basic vectors are defined as $dirs = \{d_1, d_2 \ldots d_{26}\}$ according to the pixels' relation between each other.

According to Lowe's suggestion, the next step is to rotate the coordinate axes to the direction of the key point to ensure rotation invariance. Consistent orientation is determined by fitting the differential vectors to the nearest basic vector. Each basic vector corresponds to a neighborhood block. The gradient vectors of both current block and neighborhood block

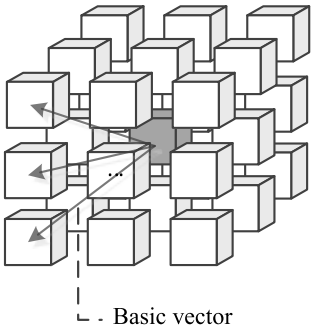

FIGURE 6. Basic vectors are determined by comparing the relationship between neighborhoods.

determine a plane, and the vertical direction of the plane is defined as the assistant orientation. After describing the consistent orientation and the assistant orientation, rotation transformation of coordinate is performed to adjust X axis and Y axis respectively to the consistent orientation and the assistant orientation. Therefore, all the vectors related to the key point can be transformed into the new coordinate system as follows:

$$V'_{n\sigma} = Trs(V_{n\sigma}) \tag{11}$$

where V′ is the output of the rotation transform. In order to effectively characterize the overall statistical distribution of the gradient vectors for the key points, the orientation vector histogram descriptor is obtained as the local feature descriptor:

$$H = \{h(d_1), h(d_2), \ldots, h(d_n) | d \in dirs\} \tag{12}$$

III. EXPERIMENTAL RESULTS

Experiments are performed on three standard pollen image datasets with a PIV computer of 2.8GHz CPU and 8GB memory. Confocal dataset is the laboratory pollen image dataset which includes 389 pollen gains in 27 different categories, in which the images were manually recorded under the confocal laser scanning microscopy in Germany [39]. All the images in the dataset take on respective appearance of different transformations especially rotation transform in order to validate the good geometric invariance of the proposed method. Pollenmonitor dataset is a real-world dataset with 22750 pollen grains from 33 categories, in which all the images were automatically collected, recorded by the first pollenmonitor prototype in Europe [40]. CHMontior dateset is a typical real-time sampling pollen image dataset, which includes 28 different categories from 23 provinces of China. Because of the irregular collection method, some pollen images are of low quality of contamination in the real collecting process. Meanwhile, all the images in three datasets were not geometrically normalized before the feature extraction. The experiments were performed by randomly taking 25% of the images of each category as training images and the remaining ones as test images. Minimum distance classifier is used for the final feature classification.

For the purpose of keeping enough local structural features and reducing the complexity of the algorithm, images are processed by using image filtering and image interpolation before feature extraction. The performance index of precision rate (PR), recall rate (RR), F1-score and recognition time (RT) are used to evaluate the recognition performance of the descriptors. F1-score is defined as:

$$F1 - score = \frac{2 \times precision \times recall}{precision + recall} \tag{13}$$

Furthermore, in order to validate the performance of the local feature descriptor proposed above, the average recognition results of the descriptor are also compared with those of the four traditional features on three datasets in three methods including SC-Zernike moment descriptors [30], STAF descriptors [26], GLCM & DTM descriptors [27] and CNN [37].

A. EXPERIMENT RESULT IN CONFOCAL DATA SET

Figure 7 presents the recognition results based on six representative pollen categories from the Confocal Dataset. It is shown that without contamination most of the Confocal pollen images that have a clear edge and abundant local texture information have been correctly classified.

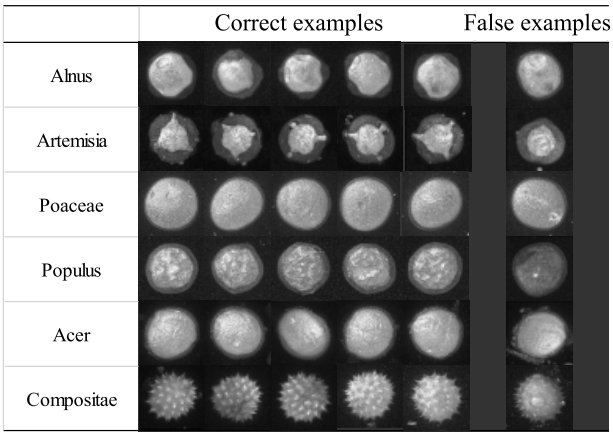

FIGURE 7. Recognition examples on 6 pollen categories from the Confocal dataset.

Table 1 shows the recognition rate of six representative pollen images. It can be inferred that the recognition performance for different pollen categories varies significantly. Among all the pollen categories, the precision rate reaches more than 93% in Compositae pollen images, which have the clearest edge and the most abundant local structure information.. As for relatively low quality pollen images, such as Poaceae, the precision rate can only reach about 76%. It can be obtained from the recognition results that the precision rate is affected by the quality of pollen images. Most of the misclassifications are mainly due to noise or uncleared sampling. The experimental results validate that the SLKP are quite suitable for practical pollen classification because of good rotation invariance.

TABLE 1. Recognition results on 6 representative pollen categories from the Confocal dataset.

	Precision (%)	Recall (%)	F1-score	RT(s)
Alnus	87.75	83.33	85.48	2.93
Chenopodium	90.25	96.62	93.33	3.48
Compositae	93.50	90.25	91.85	3.45
Poaceae	76.33	96.00	85.04	3.12
Urtica	93.33	88.85	91.03	3.01
Tilia	90.62	79.76	84.84	3.17

B. EXPERIMENT RESULT IN POLLENMONITOR DATA SET

Representative recognition examples from six pollen categories in the Pollenmonitor dataset are presented in Figure 8. The experimental results indicate that there is a great difference in image quality between Pollenmonitor dataset and Confocal dataset because of different collection methods.

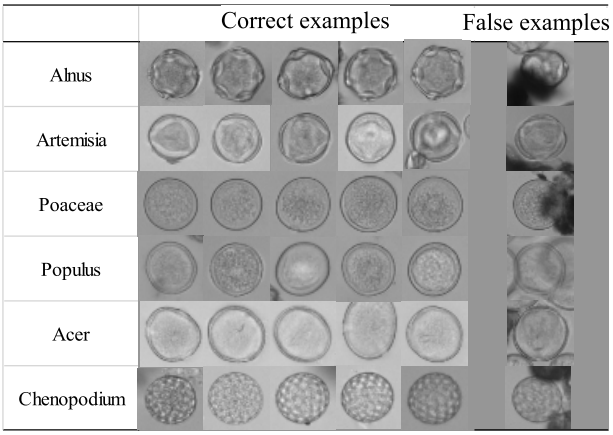

FIGURE 8. Recognition examples on 6 pollen categories from the Pollenmonitor dataset.

TABLE 2. Recognition results on 6 representative pollen categories from the POLLENMONITOR dataset.

	Precision (%)	Recall (%)	F1-score	RT(s)
Alnus	91.42	85.24	88.22	4.73
Aesculus	71.33	92.16	80.41	4.65
Tilia	79.20	84.26	81.65	4.57
Chenopodium	88.00	85.24	86.60	4.02
Populus	85.33	89.85	87.53	4.21
Corylus	87.63	75.34	81.02	4.86

Table 2 presents the detailed recognition results on the six pollen categories. It can be seen that the recognition performance also varies between categories, Although the precision rate in Pollenmonitor dataset is influenced by the quality of the pollen images, most pollen can still be classified correctly. The highest precision rate of 91.42% is obtained on the classification of the Alnus pollen, while the precision rate on the Aesculus pollen can only reach 71.33%. False examples are mainly those pollen images with low quality caused by the deformation and contamination during the collection process which affects the feature integrity of SLKP feature. Most pollen images with different posture and rotation can still be classified correctly. The experimental results further validate the effectiveness of the method.

C. EXPERIMENT RESULT IN CHMONITOR DATA SET

Representative recognition examples from six pollen categories in the CHMonitor dataset are presented in Figure 9. While the pollen images in the same category take on different appearances and structures in CHMonitor dataset and Confocal dataset, most of the pollen images from CHMonitor dataset that have clear edges and abundant local texture information can still be correctly classified.

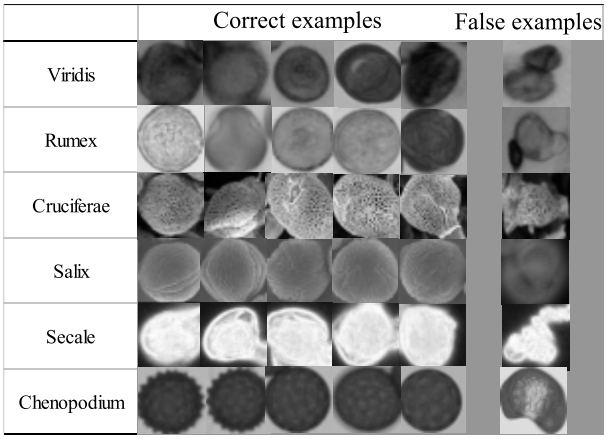

FIGURE 9. Recognition examples on 6 pollen categories from the CHMonitor dataset.

The detailed experimental results of CHMonitor are presented in Table 3. It can be inferred that recognition performances for different pollen categories vary from one another due to the differences in image quality and dimensionality of structure information. For example, among all the pollen categories, the highest precision rate of 96.35% is obtained on the classification of the Cruciferae pollen images, which have the clearest edges and the most abundant local structure information, while the precision rate on the Quercus pollen can only reach 78.84%. False examples are mainly caused by the abnormal pollen gains and obscure surface structures which may result in the variation of the extracted feature.

D. RESULTS AND DISCUSSION

The experimental results are compared with SC-Zernike descriptor, STAF descriptor, GLCM & DTM descriptor and CNN to verify the validity of the algorithm. In addition to differences of inner key points between pollen of different categories, there are different engraved textures on the external wall of the pollen grains, such as thorn, tumor, rod, cave and net, according to the variety of pollen, which may seem

TABLE 3. Recognition results on 6 representative pollen categories from the CHMONITOR dataset.

	Precision (%)	Recall (%)	F1-score	RT(s)
Alnus	89.50	87.75	88.62	2.57
Cruciferae	96.35	86.67	91.25	3.05
Salix	80.95	95.00	87.41	2.86
Chenopodium	85.62	87.46	86.53	2.25
Urtica	93.27	86.33	89.67	2.64
Quercus	78.84	100.00	88.17	2.65

more obvious in the three-dimensional pollen images. Thus, the SLKP extraction methods we proposed can present a stronger discriminating capability, better recognition results and higher recognition efficiency. Comparison results regarding the average precision rate and recognition time between the SLKP and the other three algorithms are shown in Table 4.

TABLE 4. Comparison between SLKP and other methods on average.

Method	Precision (%)	RR (%)	F1-score	RT(s/frame)
SLKP	88.25	85.56	89.46	3.76
SC-Zernike	86.32	79.85	82.96	6.90
STAF	76.45	89.36	82.40	1.98
GLCM &DTM	82.87	67.14	74.18	8.74
CNN	76.63	92.75	83.92	1.17

It's obvious that the recognition performance of the proposed SLKP is superior to that of the other three methods on the pollen images. From the experimental results, the average recognition rate of the SLKP descriptor proposed in this paper has been increased to a varying degree compared with the other three algorithms. The average identification rate for the pollen images reaches up to 88.25%, which outperforms that of the other descriptors by 11.8% compared with the STAF descriptor. Meanwhile, the time complexity of SLKP is relatively improved and the recognition time can be limited to 3.8s on average. The complexity of the algorithm is apparently reduced in the process of extracting three-dimensional features comparing with the GLCM & DTM descriptor. Besides, the average precision rate with CNN under-performs that of the proposed method by the 11.62% decline, which indicates that SLKP may have some advantages in small data set.

The corresponding classification results of average recognition precision versus average recall rates on the Confocal dataset are presented in Figure 10. It is obvious that the proposed SKLP is superior to that of the other methods on the pollen images in recognition performance. The SLKP is a lightweight statistical feature extraction method based on SIFT and histogram algorithms, which can reduce the high dimensionality of the descriptors for 3D pollen images.

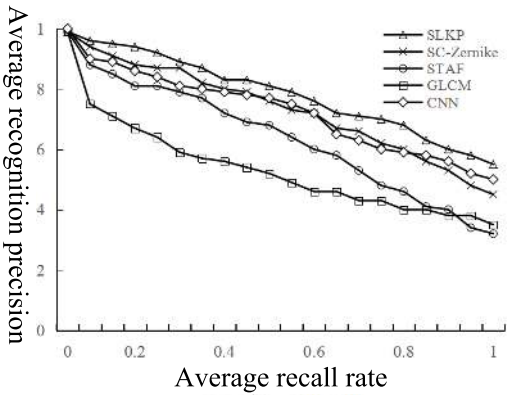

FIGURE 10. Average recognition precision rates versus average recall rates on the Confocal dataset.

Abundant textures, edges and corners of pollen images make it easier and faster for the gradient vectors to detect the stable points between layers with different scales in 3D Gaussian pyramid so that local features can be effectively extracted using the proposed method. Besides, the improvement of 3D rotation transform enhances the applicability of SLKP in the area of pollen recognition. The average recognition time, including preprocessing time and feature extraction time, is calculated to be about 3.7s on average, which is far superior to that of the other high-complexity algorithms. Satisfactory results on the pollen datasets further validate the good geometric invariance and high compactness of the proposed SLKP descriptor.

IV. CONCLUSION

This paper proposed the Spatial Local Key Point descriptor extraction method based on SIFT, and this method is applied to a pollen recognition and classification experiment on three standard pollen image datasets. It contributes to the improvement of pollen recognition in the following aspects: firstly, local key point is extracted as the original information of the image which can retain information of extrema of the image so that high dimensionality can be reduced in such method. Secondly, the experimental results show that the SLKP have great robustness to illumination, rotation, scale transformation, and affine transformation. We will hopefully contribute to improving the recognition rate and algorithm complexity in further studies. Experiments on other datasets with different quality will be performed as well.

REFERENCES

[1] R. N. Mcinnes, D. Hemming, P. Burgess, D. Lyndsay, N. J. Osborne, C. A. Skjøth, S. Thomas, and S. Vardoulakis, "Mapping allergenic pollen vegetation in UK to study environmental exposure and human health," *Sci. Total Environ.*, vols. 599–600, pp. 483–499, Dec. 2017.

[2] P. M. Leru, A. M. Eftimie, and M. Thibaudon, "First allergenic pollen monitoring in bucharest and results of three years collaboration with european aerobiology specialists," *Romanian J. Internal Med.*, vol. 56, no. 1, pp. 27–33, Mar. 2018.

[3] D. Barboni, S. P. Harrison, P. J. Bartlein, G. Jalut, M. New, I. C. Prentice, M.-F. Sanchez-Goñi, A. Spessa, B. Davis, and A. C. Stevenson, "Relationships between plant traits and climate in the Mediterranean region: A pollen data analysis," *J. Vegetation Sci.*, vol. 15, no. 5, pp. 635–646, Oct. 2010.

[4] C. H. Katelaris and P. J. Beggs, "Climate change: Allergens and allergic diseases," *Internal Med. J.*, vol. 48, no. 2, pp. 129–134, Feb. 2018.
[5] R. M. Hodgson, C. A. Holdaway, Y. Zhang, D. W. Fountain, and J. R. Flenley, "Progress towards a system for the automatic recognition of pollen using light microscope images," in *Proc. 4th Int. Symp. Image Signal Process. Anal.*, Sep. 2005, pp. 76–81.
[6] R. Navares and J. L. Aznarte, "Predicting the Poaceae pollen season: Six month-ahead forecasting and identification of relevant features," *Int. J. Biometeorol.*, vol. 61, no. 4, pp. 647–656, Apr. 2017.
[7] C. Chen, E. A. Hendriks, R. P. W. Duin, J. H. C. Reiber, P. S. Hiemstra, L. A. de Weger, and B. C. Stoel, "Feasibility study on automated recognition of allergenic pollen: Grass, birch and mugwort," *Aerobiologia*, vol. 22, no. 4, pp. 275–284, 2006.
[8] A. Daood, E. Ribeiro, and M. Bush, "Pollen recognition using a multi-layer hierarchical classifier," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 3091–3096.
[9] J. Tello, M. I. Montemayor, A. Forneck, and J. Ibáñez, "A new image-based tool for the high throughput phenotyping of pollen viability: Evaluation of inter- and intra-cultivar diversity in grapevine," *Plant Methods*, vol. 14, no. 1, p. 3, Jan. 2018.
[10] X. Bi, S. Zhu, J. Xian, L. Wei, H. Yin, M. Li, and Z. Chen, "Multihyperspectral microscopic imaging for the precise identification of pollen," *Anal. Lett.*, vol. 51, no. 14, pp. 2295–2303, Jun. 2018.
[11] Y. Xie, Z. Wang, X. Zhao, and Y. Zhu, "Orientational local binary pattern extraction method for 3D pollen image," *J. Comput.-Aided Des. Comput. Graph.*, vol. 30, no. 3, pp. 408–414, Mar. 2018.
[12] S. Kong, S. Punyasena, and C. Fowlkes, "Spatially aware dictionary learning and coding for fossil pollen identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*, Jun. 2016, pp. 1–10.
[13] R. Redondo, G. Bueno, F. Chung, R. Nava, J. V. Marcos, G. Cristóbal, T. Rodríguez, A. Gonzalez-Porto, C. Pardo, O. Déniz, and B. Escalante-Ramírez, "Pollen segmentation and feature evaluation for automatic classification in bright-field microscopy," *Comput. Electron. Agricult.*, vol. 110, pp. 56–69, Jan. 2015.
[14] A. Kubik-Komar, E. Kubera, and K. Piotrowska-Weryszko, "Selection of morphological features of pollen grains for chosen tree taxa," *Biol. Open*, vol. 7, no. 5, May 2018, Art. no. bio.031237.
[15] X. Yong-Hua, X. Zhao-Fei, and H. Burkhardt, "Spatial geometric constraints histogram descriptors based on curvature mesh graph for 3D pollen particles recognition," *Chin. Phys. B*, vol. 23, no. 6, Apr. 2014, Art. no. 060701.
[16] S. Yang, "Symbol recognition via statistical integration of pixel-level constraint histograms: A new descriptor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 278–281, Feb. 2005.
[17] J. Ren, X. Jiang, and J. Yuan, "Noise-resistant local binary pattern with an embedded error-correction mechanism," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 4049–4060, Oct. 2013.
[18] M. Reisert and H. Burkhardt, "Invariant features for 3D-data based on group integration using directional information and spherical harmonic expansion," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2006, pp. 206–209.
[19] K. A. Holt and K. D. Bennett, "Principles and methods for automated palynology," *New Phytologist*, vol. 203, no. 3, pp. 735–742, Aug. 2014.
[20] E. C. Stillman and J. R. Flenley, "The needs and prospects for automation in palynology," *Quaternary Sci. Rev.*, vol. 15, no. 1, pp. 1–5, 1996.
[21] Z. Mehmood, N. Gul, M. Altaf, T. Mahmood, T. Saba, A. Rehman, and M. T. Mahmood, "Scene search based on the adapted triangular regions and soft clustering to improve the effectiveness of the visual-bag-of-words model," *EURASIP J. Image Video Process.*, vol. 1, p. 48, Jun. 2018.
[22] Z. Mehmood, M. Rashid, A. Rehman, T. Saba, H. Dawood, and H. Dawood, "Effect of complementary visual words versus complementary features on clustering for effective content-based image search," *J. Intell. Fuzzy Syst. Preprint*, vol. 35, no. 5, pp. 5421–5434, Nov. 2018.
[23] F. Baig, Z. Mehmood, M. Rashid, M. A. Javid, A. Rehman, T. Saba, and A. Adnan, "Boosting the performance of the BoVW model using SURF-CoHOG-based sparse features with relevance feedback for cbir," *Iranian J. Sci. Technol., Trans. Elect. Eng.*, pp. 1–20, 2019.
[24] M. Ayub, M. A. Ghazanfar, Z. Mehmood, T. Saba, R. Alharbey, A. M. Munshi, and M. A. Alrige, "Modeling user rating preference behavior to improve the performance of the collaborative filtering based recommender systems," *PloS ONE*, vol. 14, no. 8, Aug. 2019, Art. no. e0220129.
[25] M. Rodriguez-Damian, E. Cernadas, A. Formella, M. Fernandez-Delgado, and P. D. Sa-Otero, "Automatic detection and classification of grains of pollen based on shape and texture," *IEEE Trans. Syst., Man, Cybern., C (Appl. Rev.)*, vol. 36, no. 4, pp. 531–542, Jul. 2006.
[26] G. Lozano-Vega, Y. Benezeth, F. Marzani, and F. Boochs, "Analysis of relevant features for pollen classification," in *Artificial Intelligence Applications and Innovations*. Berlin, Germany: Springer, 2014, pp. 395–404.
[27] J. V. Marcos, R. Nava, G. Cristóbal, R. Redondo, B. Escalante-Ramírez, G. Bueno, Ó. Déniz, A. González-Porto, C. Pardo, F. Chung, and T. Rodríguez, "Automated pollen identification using microscopic imaging and texture analysis," *Micron*, vol. 68, no. 9, pp. 36–46, Jan. 2015.
[28] O. Ronneberger, H. Burkhardt, and E. Schultz, "General-purpose object recognition in 3D volume data sets using gray-scale invariants—Classification of airborne pollen-grains recorded with a confocal laser scanning microscope," in *Proc. Object Recognit. Supported User Interact. Service Robots*, vol. 2, Aug. 2002, pp. 290–295.
[29] G. Kim, S. Lee, S. Shin, and Y. Park, "Three-dimensional label-free imaging and analysis of Pinus pollen grains using optical diffraction tomography," *Sci. Rep.*, vol. 8, no. 1, p. 1782, Jan. 2018.
[30] Y. Xie and Z. Xu, "A Fourier descriptor based on Zernike invariant moments in spherical coordinates for 3D pollen image recognition," in *Proc. 8th Int. Conf. Biomed. Eng. Inform. (BMEI)*, Oct. 2015, pp. 453–457.
[31] O. Ronneberger, Q. Wang, and H. Burkhardt, "3D invariants with high robustness to local deformations for automated pollen recognition," in *Proc. Joint Pattern Recognit. Symp.*, 2007, pp. 425–435.
[32] A. Sarwar, Z. Mehmood, T. Saba, K. A. Qazi, A. Adnan, and H. Jamal, "A novel method for content-based image retrieval to improve the effectiveness of the bag-of-words model using a support vector machine," *J. Inf. Sci.*, vol. 45, no. 1, pp. 117–135, 2019.
[33] K. Qazi, T. Nawaz, Z. Mehmood, M. Rashid, and H. A. Habib, "A hybrid technique for speech segregation and classification using a sophisticated deep neural network," *PloS ONE*, vol. 13, Mar. 2018, Art. no. e0194151.
[34] U. Sharif, Z. Mehmood, T. Mahmood, M. A. Javid, A. Rehman, and T. Saba, "Scene analysis and search using local features and support vector machine for effective content-based image retrieval," *Artif. Intell. Rev.*, vol. 52, no. 2, pp. 901–925, Aug. 2019.
[35] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
[36] A. Daood, E. Ribeiro, and M. Bush, "Sequential recognition of pollen grain Z-stacks by combining CNN and RNN," in *Proc. 21st Int. Flairs Conf. (FLAIRS)*, May 2018.
[37] N. Khanzhina, E. Putin, A. Filchenkov, and E. Zamyatina, "Pollen grain recognition using convolutional neural network," in *Proc. Eur. Symp. Artif. Neural Netw. (ESANN)*, 2018, pp. 1–6.
[38] T. Sledevi , "The application of convolutional neural network for pollen bearing bee classification," in *Proc. IEEE 6th Workshop Adv. Inf., Electron. Elect. Eng.*, Nov. 2018, pp. 1–4.
[39] O. Ronneberger, E. Schultz, and H. Burkhardt, "Automated pollen recognition using 3D volume images from fluorescence microscopy," *Aerobiologia*, vol. 18, no. 2, pp. 107–115, Jun. 2002.
[40] M. Ranzato, P. E. Taylor, J. M. House, R. C. Flagan, Y. LeCun, and P. Perona, "Automatic recognition of biological particles in microscopic images," *Pattern Recognit. Lett.*, vol. 28, no. 1, pp. 31–39, Jan. 2007.

ZHUO WANG received the master's degree in computer science from the Nanjing University of Information Science and Technology, in 2018. He is currently with the School of Information and Electrical Engineering, Xuzhou University of Technology, Xuzhou, China. His research interests include feature extraction, pattern recognition, and machine learning.

WENZHENG BAO received the master's degree in computer science from the University of Jinan, in 2015, and the Ph.D. degree in computer science from Tongji University, in 2018. He is currently with the School of Information and Electrical Engineering, Xuzhou University of Technology, Xuzhou, China. His research interests include bioinformatics and machine learning.

DA LIN received the master's degree in computer science from Jilin University, in 2007, and the Ph.D. degree in biomedical engineering from the Dalian University of Technology, in 2011. He is currently with the School of Information and Electrical Engineering, Xuzhou University of Technology, Xuzhou, China. His research interests include intelligent control, image processing, and machine learning.

ZIXUAN WANG was born in 1994. She graduated from Sun Yat-sen University, in 2012, and received the master's degree in management (finance) from Durham University, in 2018. Her current interests include machine learning and data envelopement analysis.