Classifying Pollen Using Robust Sequence Alignment of Sparse Z-Stack Volumes

Amar Daood[1(✉)], Eraldo Ribeiro[2], and Mark Bush[3]

[1] Department of Electrical and Computer Engineering, Florida Institute of Technology Melbourne, Melbourne, FL, USA
adaood2012@my.fit.edu

[2] Department of Computer Sciences and Cybersecurity, Florida Institute of Technology Melbourne, Melbourne, FL, USA
eribeiro@fit.edu

[3] Department of Biological Sciences, Florida Institute of Technology Melbourne, Melbourne, FL, USA

Abstract. The identification of pollen grains is a task needed in many scientific and industrial applications, ranging from climate research to petroleum exploration. It is also a time-consuming task. To produce data, pollen experts spend hours, sometimes months, visually counting thousands of pollen grains from hundreds of images acquired by microscopes. Most current automation of pollen identification rely on single-focus images. While this type of image contains characteristic texture and shape, it lacks information about how these visual cues vary across the grain's surface. In this paper, we propose a method that recognizes pollen species from stacks of multi-focal images. Here, each pollen grain is represented by a multi-focal stack. Our method matches unknown stacks to pre-learned ones using the Longest-Common Sub-Sequence (LCSS) algorithm. The matching process relies on the variations of visual texture and contour that occur along the image stack, which are captured by a low-rank and sparse decomposition technique. We tested our method on 392 image stacks from 10 species of pollen grains. The proposed method achieves a remarkable recognition rate of 99.23%.

1 Introduction

The classification of pollen grains is the main data-collection task in many disciplines including ecology, forensic sciences, allergy control, and oil exploration [1]. In most of these applications, the task of counting and identifying pollen grains require pollen experts (i.e., palynologists) to spend hours looking at pollen under the microscope. Automation of pollen identification can reduce data-collection times from months to a few hours [2].

Since Flenley [2] suggested the automation of pollen identification as a means to increase research throughput in Palynology, a number of identification methods have been proposed to describe the visual appearance of pollen grains. For example, shape features such as roundness, perimeter, and area have been used

G. Bebis et al. (Eds.): ISVC 2016, Part I, LNCS 10072, pp. 331–340, 2016.
DOI: 10.1007/978-3-319-50835-1_31

by [1,3]. Texture measurements such as gray-level co-occurrence matrix, Gabor features, and moment invariants were used by [4–6]. These methods work on images of pollen grains acquired under a single focal plane.

Instead of using isolated visual characteristics of a single focal plane, some methods combined multiple visual features extracted from multiple focal z-planes. Chica [7] identified five pollen classes by using shape and texture features from three focal images. The shape features were area, perimeter, diameter, and the texture features were mean, standard deviation, and the entropy of gray-level histograms. Lagerstrom et al. [8] extracted histogram statistics, moments, grey-level co-occurrence matrix, and Gabor features from nine focal planes to classify 15 types of pollen grains. Shape and texture features were also combined in [9] using gray-level co-occurrence matrix, Gabor features, local binary patterns, and moments to classify 15 pollen types.

Although the combination of visual features works well for pollen identification, the use of stacked multi-focal images is still underexploited. In this paper, we propose to classify pollen grains using multi-focal image sequences. We want to capture characteristic information of changes in visual appearance that occur on the grain's surface across multiple focal planes (i.e., z-stack). Our method commences by processing the entire z-stack volume using a low-rank and sparse decomposition [10] to extract the visual changes across focal planes. These changes will appear in the sparse component of the z-stack. From this sparse volume, we extract features from each one of its planes. Our feature-extraction method further decomposes each slice of the sparse volume into multiple concentric regions by clustering the gray-level intensity and their associated polar coordinates. Shape and texture features are then extracted from each layer. These features are used by a nearest-neighbor classifier that employs the Longest-Common Sub-Sequence (LCSS) algorithm to match the multifocal appearance descriptors. We tested our method on a dataset of 10 pollen types. The dataset has 392 z-stack sequences with 10 focal-planes for each sample. Figure 1 shows one sample z-stack for each pollen type in our dataset.

2 Method

Our method begins by processing the entire z-stack volume of the pollen grain using a low-rank and sparse decomposition. Then, we process each focal image inside the sparse volume to create multiple layers to represent the visual appearance of the pollen. Finally, we classify the pollen grain types using a sequence-alignment method which considers the multifocal volume as an image sequence. The sequence-matching algorithm computes the similarity between the appearance model of the entire volume for two pollen grains. Figure 2 shows a block diagram of the proposed method.

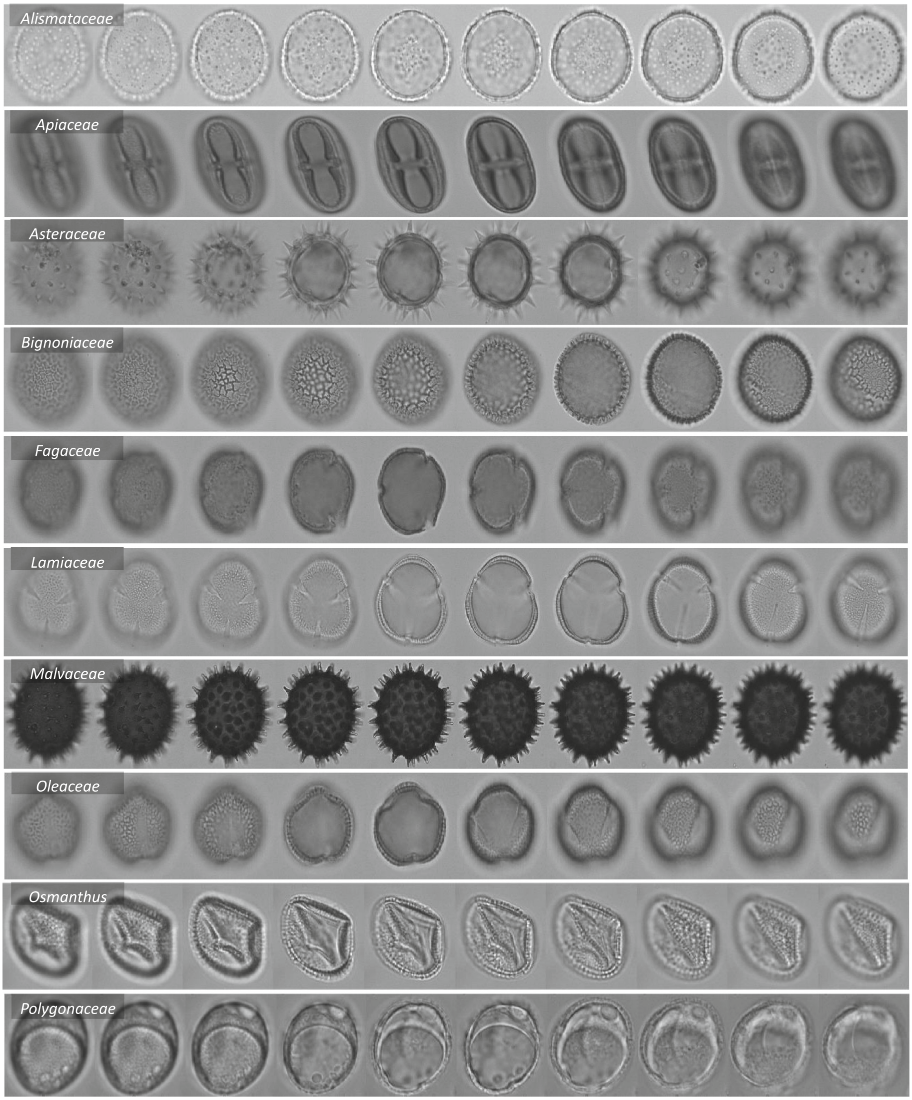


Fig. 1. Z-stacks from pollen types in our dataset. Dataset has 392 stacks of ten images of 10 pollen types. Images were provided by the Florida Tech's Paleoecology Laboratory.

2.1 Low Rank and Sparse Decomposition

The low-rank and sparse decomposition decomposes a matrix into two components: low-rank and sparse components. Given a data matrix $A \in R^{m \times n}$, it assumes the matrix A can be decomposed as:

$$A = L + S, \tag{1}$$

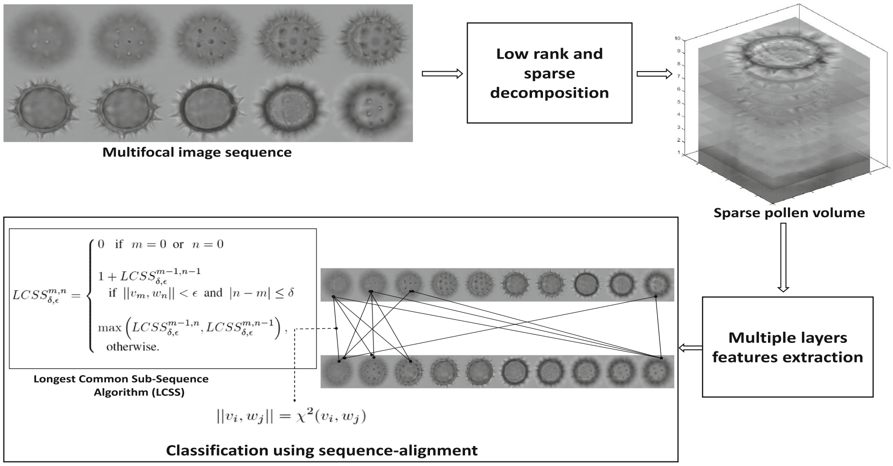


Fig. 2. Method overview. It takes multifocal image sequence as input. The sequence is processed using low-rank-sparse decomposition to create a sparse volume. Each focal plane of the sparse volume is represented by multiple features to create a sequence of appearance models. Finally, this sequence is matched to the known ones in a database using the LCSS alignment algorithm.

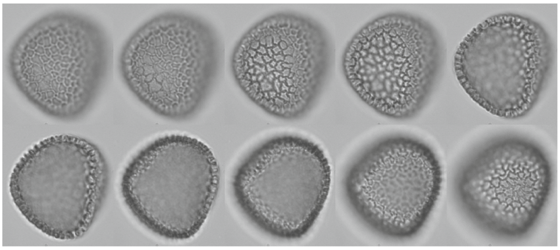
where $L \in R^{m \times n}$ is low-rank matrix and $S \in R^{m \times n}$ is the sparse matrix. This decomposition has been used widely to process sequential data such as videos. For pollen classification, we use the low-rank and sparse analysis on the z-stack to find regions of interest and decompose the entire volume to stationary and non-stationary pixels. We assume that changes in focus occurs only in regions of the pollen's surface, and that the information of the background and any noise is static inside the multifocal sequence. Our method uses the LRS library [11] that implements the robust principal component analysis (RPCA) [10]. RPCA solves the following optimization problem:

$$\min_{L,S} \; rank(L) + \lambda \|S\|_0 \quad \text{subject to} \;\; L + S = A, \tag{2}$$

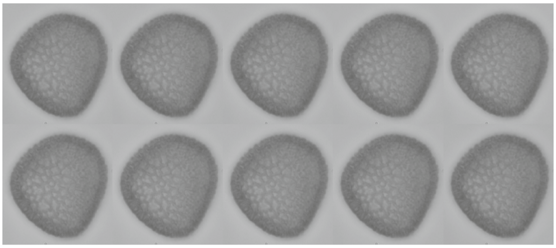
where A is the z-stack matrix. Parameter λ controls the trade-off between sparsity and the low rank. A decomposition example is shown in Fig. 3.

As shown in Fig. 3, low rank and sparse decomposition separate z-stack pixels into two components. The low-rank component combines all z-stack frames into one image (Fig. 3b), essentially compressing the entire volume to a single focused plane. In contrast, the sparse component emphasizes the gray-level changes across the entire volume. The sparse volume helps recognition of the multi-focal stacks because it describes the changes across the stack planes. We use the sparse component to describe the characteristics of the pollen z-stacks.

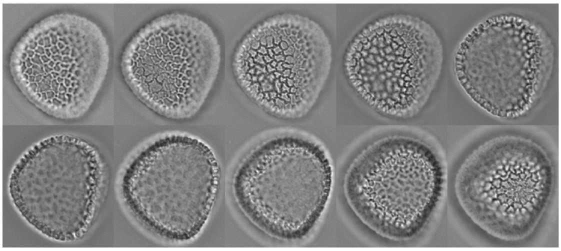
Prior to extracting features, we apply low-rank-sparse decomposition to all z-stacks. With the sparse volumes at hand, we extract features from each individual focal plane. The feature-extraction process is described next.



(a) The original sequence



(b) Low-rank component



(c) Sparse component

Fig. 3. Decomposing a pollen grain z-stack using RPCA.

2.2 Feature Extraction

We adopt a version of our previous work [12] to extract features from multiple concentric regions from pollen images. First, we normalize the image contrast using histogram equalization. Then, we decompose the image into multiple layers using clustering where each cluster represents a layer of pollen grain regions. The final set of regions is given by:

$$R = \{L_1, L_2, \ldots, L_d\}\,, \tag{3}$$

where d is the number of layers and L represents an individual layer of a image. We combine spatial information and intensity to perform the clustering process. Polar mapping is used to flattens the near-circular pollen region into a rectangular shape (Fig. 4). Each segmented pollen is represented by a distance-angle map (r, θ), where r is the distance from boundary pixels to the grain's estimated centroid, and θ is the angle of vector r. After that, we combine the intensity and the polar coordinates to create vector $v = (r, \theta, i)$. Then, we cluster the pollen image into multiple regions where each layer is given by:

$$L_i = \{c_r, c_\theta, c_i, V_{r,\theta,i}\}, \tag{4}$$

where $V_{r,\theta,i}$ contains pixels and polar information of each layer. To keep the layers ordering, we reorder the layers according to the center of each layer c_r. This sorting process keeps the layer ordering consistent from outer layers to inner layers. Finally, we reverse the polar mapping to obtain the original Cartesian coordinates. After that, feature extraction process is performed on each layer. Extracting features for each layer individually improves the representation of the visual information. We adopt different types of features. The local binary pattern histogram and fractal dimension are used to describe each layer. In addition, gray level and histogram statistics are extracted and combined to create feature vectors. Figure 5 summarizes the method.

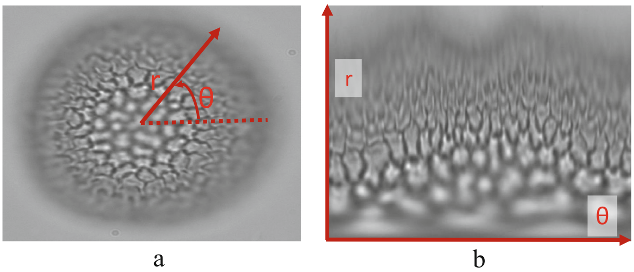


Fig. 4. Polar mapping. (a) Original image. (b) Transformed image.

2.3 Recognition Using Sequence Matching

We use a sequence-alignment method to measure the similarity between two sequences. This scheme considers the appearance variations in the sequence of multifocal pollen images rather than classifying the concatenated features blindly. We use the Longest Common Sub sequence (LCSS) matching scheme introduced in [13]. We define a distance measure between two sequences. Let $V = (v_1, \ldots, v_m)$ and $W = (w_1, \ldots, w_n)$ represent two multifocal image sequences for two different pollen grains. The measure used in the LCSS scheme can be defined for an integer δ and a real number $0 < \epsilon < 1$ as follows:

$$LCSS_{\delta,\epsilon}^{m,n} = \begin{cases} 0 \quad \text{if} \quad m = 0 \ \text{or} \ n = 0 \\ \\ 1 + LCSS_{\delta,\epsilon}^{m-1,n-1} \\ \quad \text{if} \ \|v_m, w_n\| < \epsilon \ \text{and} \ |n - m| \leq \delta \\ \\ \max\left(LCSS_{\delta,\epsilon}^{m-1,n}, LCSS_{\delta,\epsilon}^{m,n-1}\right), \\ \quad \text{otherwise.} \end{cases} \tag{5}$$

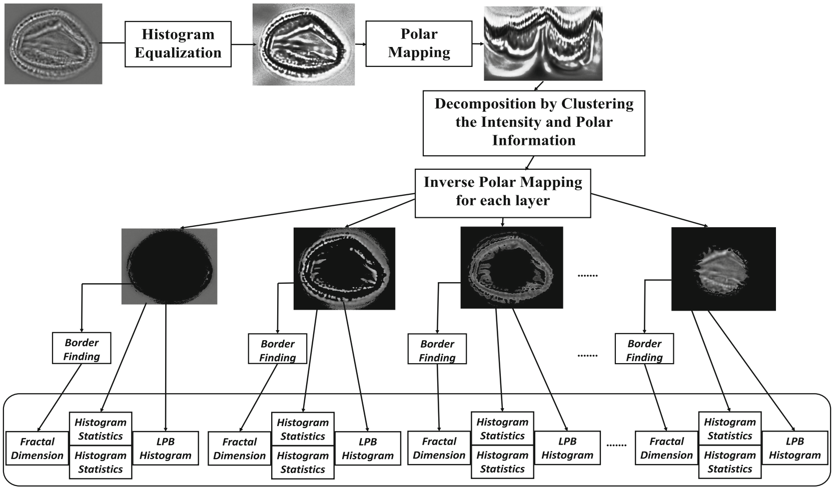


Fig. 5. Feature-extraction method. First, we use histogram equalization to enhance image contrast and then we apply a polar mapping to capture spatial information. A clustering process decomposes the image into concentric layers. Finally, we reserve the polar coordinates to obtain the original Cartesian coordinates of the layers. Feature vectors are extracted from each layer.

Here, $LCSS^{i,j}_{\delta,\epsilon} = LCSS_{\delta,\epsilon}\left((v_1, \ldots, v_i), (w_1, \ldots, w_j)\right)$, and $||.,.||$ is the distance between points v_m and w_n. The threshold ϵ controls the maximum distance between a pair of matched points, and δ controls the maximum number of consecutive unmatched points. The similarity between two sequences is given by:

$$D_{\delta,\epsilon}(V, W) = 1 - \frac{LCSS^{m,n}_{\delta,\epsilon}}{\min(m, n)}. \tag{6}$$

The distance between two focal slices as required by Eq. 5 is defined as:

$$||v_i, w_j|| = \chi^2(v_i, w_j), \tag{7}$$

where χ^2 is the chi-square similarity. We create multifocal visual models for a set of known pollen grains. These models (templates) are stored to be compared with unknown sequences for identification. Formally, given templates $\{\widehat{V}_1, \ldots, \widehat{V}_M\}$, recognition of an unknown sequence V is done by a nearest-neighbor classifier based on the similarity measure produced by the LCSS algorithm.

3 Results

We evaluated the effectiveness of the proposed method using two experiments. First, we used a leave-one-out classification technique then we divided our dataset into 75% as a training and 25% as a testing set.

We began by applying a leave-one-out classification technique. We selected one sequence as a test sample and considered the rest of our data for training. Then, we classified the test sample according to the training templates using a nearest-neighbor classifier. In this case, we achieved about 99.23% recognition rate. Then, we divided our dataset into 75% as a training and 25% as a testing set to compare with the traditional techniques that extracted and concatenated feature directly. We extracted the following features: histogram features (i.e., mean and variance of histogram), gray-level statistics (i.e., mean, variance, and entropy), geometrical features (i.e., area, perimeter, compactness, roundness, and aspect ratio based on minor and major axises), fractal dimension, gray level co-occurrence matrix (GLCM), moments invariant, Gabor features, histograms of oriented gradient (HOG) descriptors, and local binary pattern histogram (LBP). Then, a support vector machine classifier was trained using these features. Additionally, we reproduced the results of two works in the literature that used concatenated features from multifocal planes: Chica's Method [7] and Lagerstrom's Method [8]. The obtained results are shown in Table 1.

Table 1. Classification rates

Method	Classification (%)
Histogram features, Gray level statistics	81.92%
Geometrical features, fractal dimension	80.12%
Gray level co-occurrence matrix	73.44%
Moments invariants	70.35%
Gabor features	76.04%
HOG	75.63%
LBP	84.73%
Chica's Method	86.18%
Lagerstrom's Method	83.96%
Histogram, gray-level statistics, fractal dimension, LBP	88.88%
Our proposed method	98.66%

To show the superiority of our method over traditional approaches, we compared the proposed method with a feature-combination method that combined histogram, gray-level statistics, fractal dimension, and LBP as features. This method achieved a 88.88% classification rate. After we applied a significance test, the P-value was 9.76×10^{-4} which rejected the null hypothesis. Furthermore, we computed some classification metrics in Table 2 including the average of precision, recall, sensitivity, specificity, and F-score.

Table 2. Evaluation measurements

Method	Precision	Recall	Sensitivity	Specificity	F score
Features combination	89.77%	89.07%	89.07%	98.75%	88.98%
Our method	**99.09%**	**99.00%**	**99.00%**	**99.89%**	**99.03%**

4 Conclusion and Future Work

We proposed a recognition algorithm to identify pollen grains using multifocal sequences of images. The main novelty in our algorithm is twofold. Firstly, we decomposed the pollen slice to create multiple layers by clustering intensity and polar information as an extra stage before performing features extraction. We captured the visual information of each slice by combining shape and texture features from each layer to build the appearance model. Secondly, instead of concatenating features from different focal-planes, we identified the pollen samples using sequence alignment technique to consider the sequence's effect in our recognition process. Additionally, our method does not require any specific set up of a focal length and thus it is applicable for other microscopic objects based on multifocal sequences. For future work, we plan to test our method on a larger number of species. We also plan to use the low-rank volume as a prior to pre-classify pollen into subgroups.

References

1. Treloar, W.J., Taylor, G.E., Flenley, J.R.: Towards automation of palynology 1: analysis of pollen shape and ornamentation using simple geometric measures, derived from scanning electron microscope images. Quat. Sci. **19**, 745–754 (2004)
2. Flenley, J.: The problem of pollen recognition. In: Clowes M.B., Penny J.P. (eds.) Problems in Picture Interpretation, pp. 141–145. CSIRO (1968)
3. Boucher, A., Hidalgo, P., Thonnat, M., Belmonte, J., Galan, C., Bonton, P., Tomczak, R.: Development of a semi-automatic system for pollen recognition. Aerobiologia **18**, 195–201 (2002)
4. Fernandez-Delgado, M., Carrion, P., Cernadas, E., Galvez, J.F.: Improved classification of pollen texture images using SVM and MLP. In: 3rd IASTED International Conference Visualization, Imaging and Image Processing (2003)
5. Zhang, Y., Fountain, D., Hodgson, R., Flenley, J., Gunetileke, S.: Towards automation of palynology 3: pollen pattern recognition using Gabor transforms and digital moments. Quat. Sci. **19**(8), 763–768 (2004)
6. Li, P., Treloar, W.J., Flenley, J.R., Empson, L.: Towards automation of palynology 2: the use of texture measures and neural network analysis for automated identification of optical images of pollen grains. Quat. Sci. **19**, 755–762 (2004)
7. Chica, M.: Authentication of bee pollen grains in bright-field microscopy by combining one-class classification techniques and image processing. Microsc. Res. Tech. **75**, 1475–1485 (2012)
8. Lagerstrom, R., Arzhaeva, Y., Bischof, L., Haberle, S., Hopf, F., Lovell, D.: A comparison of classification algorithms within the classifynder pollen imaging system. In: International symposium on CMLS (2013)

9. Marcos, J.V., Nava, R., Cristobal, G., Redondo, R., Escalante-Ramirez, B., Bueno, G., Deniz, O., Gonzalez-Porto, A., Pardo, C., Chung, F., Rodriguez, T.: Automated pollen identification using microscopic imaging and texture analysis. Micron **68**, 36–46 (2015)
10. Wright, J., Ganesh, A., Rao, S., Peng, Y., Ma, Y.: Robust principal component analysis: exact recovery of corrupted low-rank matrices via convex optimization. In: NIPS, pp. 2080–2088 (2009)
11. Sobral, A., Bouwmans, T., Zahzah, E.h.: LRSlibrary: low-rank and sparse tools for background modeling and subtraction in videos. In: Robust Low-Rank and Sparse Matrix Decomposition: Applications in Image and Video Processing (2015)
12. Daood, A., Ribeiro, E., Bush, M.: Pollen grain recognition based on a multi-layer feature decomposition. In: FLAIRS-29 (2016)
13. Vlachos, M., Hadjieleftheriou, M., Gunopulos, D., Keogh, E.: Indexing multi-dimensional time-series with support for multiple distance measures. In: ACM SIGKDD, pp. 216–225 (2003)