Spatially Aware Dictionary Learning and Coding for Fossil Pollen Identification

Fangfang Li

June 6, 2018

Abstract

We propose a robust approach for performing automatic species-level recognition of fossil pollen grains in microscopy images that exploits both global shape and local texture characteristics in a patch-based matching methodology. We introduce a novel criteria for selecting meaningful and discriminative exemplar patches. We optimize this function during training using a greedy submodular function optimization framework that gives a nearoptimal solution with bounded approximation error. We use these selected exemplars as a dictionary basis and propose a spatiallyaware sparse coding method to match testing images for identification while maintaining global shape correspondence. To accelerate the coding process for fast matching, we introduce a relaxed form that uses spatiallyaware softthresholding during coding. Finally, we carry out an experimental study that demonstrates the effective-ness and efficiency of our exemplar selection and classification mechanisms, achieving 86.13 accuracy on a difficult finegrained species classification task distinguishing three types of fossil spruce pollen.

1. Introduction

As one of the most ubiquitous of terrestrial fossils, pollen has an extraordinarily rich record and has been used to test hypotheses from a broad cross-section of biological and geological sciences and a diverse array of disciplines. Detecting and classifying pollen grains in a collected sample allows one to estimate the diversity of plant species in a particular area, carry out paleoecological and paleoclimatological investigations across hundreds to millions of years, implement the identification of plant speciation and extinction events, calculate the correlation and biostratigraphic dating of rock sequences, and conduct studies of long-term anthropogenic impacts on plant communities and the study of plantpollinator relationships.

While highthroughput microscopic imaging allows for ready acquisition of large numbers of images of modern or fossilized pollen samples, identifying and counting by eye the number of grains of each species is painstaking work and requires substantial expertise and training. In this paper, we tackle the problem of performing automated species-level classification of individual pollen grains using machine learning techniques based on sparse coding to capture fine-grained distinctions in surface texture and shape. A number of previous works have proposed to apply machine learning to pollen identification. However, all of these methods have largely avoided the difficult problem of specieslevel classification, which is significant to the reconstruction of paleoenvironments and discrimination of paleoecologically and paleoclimatically significant taxa. Recently, Punyasena et al. proposed two different machine learning-based approaches to identify two pollen species of spruce. Their approach uses three categories of handcrafted features, including intensity distribution, gross shape, and texture features which are further enriched by varying the parameters for each feature computation. They show effectiveness of the approach in identifying both modern and fossil pollen grains as a threeway classification problem, and suggest that the pollen grain size and texture are important variables in pollen species discrimination. However, they rely on leaveone out validation to estimate performance and leave open the question of generalization to held out test data and other species. In this paper, we propose a robust framework to automatically identify the species of fossil pollen grains in microscopy images. There are several difficulties that arise, including the arbitrary viewpoint of the pollen grains imaged, as shown in Figure 1 and very limited amounts of expert-labeled training data (relative to many modern computer vision tasks). To address these problems, we introduce an exemplar matching strategy for identification based on local surface patches through several novel technical components. First, we propose a greedy method for selecting discriminative exemplar patches based on optimizing a submodular set function. We show our greedy algorithm is efficient and gives a near-optimal solution with a eapproximation bound. Second, we use the selected exemplar patches as a codebook dictionary and propose a spatially aware sparse coding method to match test image patches for classification. Finally, to accelerate the match-
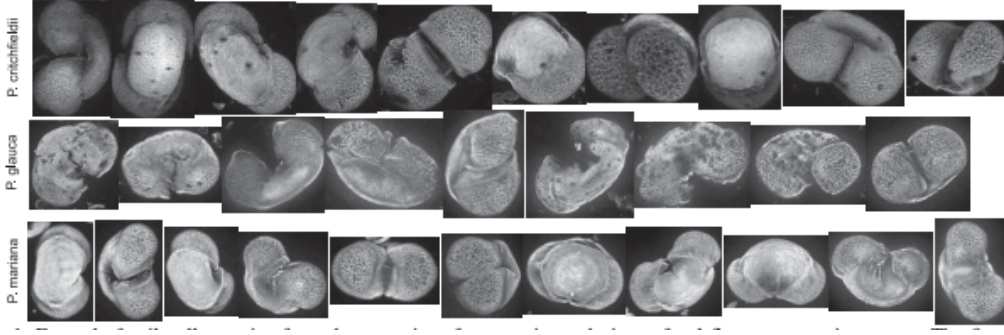
Figure 1. Example fossil pollen grains from three species of spruce, imaged via confocal fluorescence microscopy. The fine-grained identification of pollen species is not a trivial task and depends on subtle differences in the overall pollen grain shape as well as local surface texture. The arbitrary viewpoint, substantial intra-species shape variance and sample degradation of the grains poses further difficulties.

ing process for classification, we introduce a relaxed form of our weighted sparse coding method for fast matching. Through experimental study on a dataset of spruce pollen grains, we demonstrate the efficiency and effectiveness of our patch selection and classification mechanisms. Our method achieves 86.13accuracy on a threeway classification task which is quite promising given the visual difficulty of the task and the small training set size.

2. Discriminative Patch Selection

In order to allow robust matching of surface texture and local shape features of pollen grains while maintaining invariance to arbitrary viewpoint, we use a patchbased representation of appearance. Our first step is to select a small number of exemplar patches from the training dataset [2]. The selected patches or exemplars should not only represent the pollen grains well in the feature space, but also have the capability to distinguish species-level characteristics by preserving the spatial structure of the grains. We use the selected patches as a dictionary basis to match testing images for identification [1]. To this end, we formulate an objective function that scores a set of candidate patches selected from the training set based on several criteria including representational and discriminative power, and balanced sampling across classes and spatial locations. Selecting a subset of patches that optimizes this objective reduces to a well studied problem of maximizing a submodular set function, which we introduce briefly before describing the specific terms in our patch selection objective function [3]. We generate a large set of candidate patches by sampling randomly and uniformly over spatial locations across the collection of training images. The patches could be represented by pixel values or other features such as SIFT. In our experiments, we use activations from a pretrained CNN as our feature descriptor. We assume the subset of selected patch-

es should be representative of all the patches in the feature space and yield discriminative compact clusters that are balanced across classes. In addition patches should be spatially cover most regions of the pollen grain. We now describe terms that encode each of these criteria.

2.1. Patch Selection Objective Function

We generate a large set of candidate patches by sampling randomly and uniformly over spatial locations across the collection of training images. The patches could be represented by pixel values or other features such as SIFT. In our experiments, we use activations from a pretrained CNN as our feature descriptor. We assume the subset of selected patches should be representative of all the patches in the feature space and yield discriminative compact clusters that are balanced across classes. In addition patches should be spatially cover most regions of the pollen grain. We now describe terms that encode each of these criteria. We define the score of a set example A as: The process of neural networks can be expressed as:

$$\mathcal{F}_R = \sum_{j \in v} \max_{i \in A} S_{ij} \qquad (1)$$

This function is a monotonically increasing submodular function and can be seen as a special case of the facility location problem where the costs of all the nodes are the same.

References

[1] R. E. Fikes and N. J. Nilsson. Owards automation of palynology 1: Analysis of pollen shape and ornamentation using simple geometric measures, derived from scanning electron microscope images. *Journal of Quaternary Science*, 19(8):189–208, 1971. 2

[2] B. K. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 11(1):185–203, 1981. 2

[3] S. Kong, S. Punyasena, and C. Fowlkes. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pages 260–265, 2012. 2