Image-based detection and classification of allergenic pollen

Gildardo Lozano Vega

SPIM

Thèse de Doctorat

école doctorale **sciences pour l'ingénieur et microtechniques**

UNIVERSITÉ DE BOURGOGNE

Image-based Detection and Classification of Allergenic Pollen

GILDARDO LOZANO VEGA

SPIM

Thèse de Doctorat

école doctorale **sciences pour l'ingénieur et microtechniques**

UNIVERSITÉ DE BOURGOGNE

N° | X | X | X |

THÈSE présentée par

GILDARDO LOZANO VEGA

pour obtenir le

Grade de Docteur de
l'Université de Bourgogne

Spécialité : **Instrumentation et Informatique de l'Image**

Image-based Detection and Classification of Allergenic Pollen

Unité de Recherche :
Laboratoire le2i

Soutenue publiquement le 18 juin 2015 devant le Jury composé de :

BRUNO EMILE	Rapporteur	MCF-HDR à l'Université d'Orléans, Châteauroux, France
JOACHIM OHSER	Rapporteur	Professeur à la Hochschule Darmstadt, Allemagne
RONALD RÖSCH	Examinateur	Docteur au Fraunhofer ITWM Kaiserlautern, Allemagne
CHRISTIAN DAUL	Examinateur	Professeur à l'Université de Lorraine, France
FRANCK MARZANI	Co-Directeur	Professeur à l'Université de Bourgogne, France
FRANK BOOCHS	Co-Directeur	Professeur à la Hochschule Mainz, Allemagne
YANNICK BENEZETH	Co-Encadrant	MCF à l'Université de Bourgogne, France

ACKNOWLEDGEMENTS

I would like to thank first to my supervisors Prof. Franck Marzani, Prof. Frank Boochs and MCF Yannick Benezeth for their guidance and advice throughout the develepment of the thesis. Their invaluable observations were indispensable for the correct conclusion of this work. Also my gratitude is to the Institute for Spatial Information and Surveying Technology (i3mainz), the Laboratory of Electronics, Informatics and Image (Le2i), their staff and colleagues for hosting me during my PhD period, procuring the adequate resources and facilities, and for their advice and support. In particular I thank to Matthias Uhler and Celeste Chudyk for their direct contributions and discussions which led to improve my research.

I am deeply grateful to the members of the Department of Multiphase Chemistry at the Max Planck Institute for Chemistry in Mainz, which voluntarily permitted the use of their laboratory and their resources for the creation of the pollen slides, and to Celeste for her enormous dedication at the microscope. Many thanks to Dr. Reinhard Wachter for supporting me voluntarily with his wide experience on palynology.

I greatly appreciate the review and very interesting comments that the jury members of my thesis committee made to help me to improve this manuscript.

The encouragement of my family and friends was crucial to the completion of this period. Thanks to my father, the great Hernández family, Elodie, Helmut, Monika, Tobi, and Jorge.

A very special gratitude is to my two-member family for her infinite encouragement and permanent trust. Karla has been experiencing around the clock all the positive and negative effects of this adventure, and still never stepped down. Such sort of love makes great things happen.

Contents

1

Introduction

1.1/ Context and Motivation

The prevalence of allergic rhinitis is considerably high, affecting in conservative estimations 8.0%, and up to 25% of the world population [Bousquet, J. et al., 2008; Dykewicz and Hamilos, 2010]. Allergic rhinitis causes symptoms like sneezing, nasal itching, nasal congestion, rhinorrhea and in some cases eyelid swelling, reducing the quality of life of affected subjects. Moreover, this disorder is associated with the exacerbation of asthma [Bousquet, J. et al., 2008]. Allergic rhinitis is triggered by the inhalation of particles carrying allergens, like insect residues, animal dander, mold, spores, or pollen.

Being pollen one of the most important outdoor carriers, its study has become very important for the combat of the disease. The complexity of this study lies in the fact that pollen taxa exhibit different groups (i.e. types) of allergens. Therefore, the allergenic severity varies from taxon to taxon [Negrini, 1992]. The knowledge of the particular concentration of the pollen taxa to which patients are exposed is critical for the correct diagnosis and customized treatment.

The estimation of the daily pollen concentration is employed to issue alerts of high concentration to the local population and to create forecasting models. A difficulty for this analysis is the intermittence of the pollen presence in the environment due to the flowering period, the weather, pollution, climate change, and regional differences in vegetation [D'Amato et al., 2007].

Traditional methods of concentration estimation consist in counting manually one by one airborne pollen particles after being gathered by specialized devices (e.g. the Hirst volumetric sampler). The arduous work of analyzing particles on a slide under the microscope requires a considerable investment of time, especially in the high flowering season, and the specialized labor of palynologists, who spend many hours in training.

Other applications rely also on the recognition of pollen taxa. For example, paleontological reconstruction of vegetation employs similar manual counting [Jackson et al., 2000]. Likewise, the forensic science utilizes the identification of pollen particle in order to identify links between evidence, people and places [Mildenhall et al., 2006].

Time-consuming and costly labor is the main factor that limits considerably the frequency and the extent of the region to be analyzed. A clear example is Germany, where there are only 44 pollen collection stations, leaving large regions without the service of a close station. When information is required in such places, the pollen concentration is inaccurately estimated by the extrapolation of closest stations.

Additionally, manual counting is highly susceptible to human error and inconsistency. Experiments on taxonomical categorization of plankton conducted by Culverhouse *et al.* indicated that 20% of the analyses presented a variation greater than 10% [Culverhouse et al., 2014]. They indicate that classification experts can achieve only up to 80% of self-consistency in repeated analysis. The authors suggest that among the human factor that contributes to this variation are fatigue, boredom and prior expectations. They indicate firmly that this variation can be found on many others ecological studies.

All these aforementioned factors impact largely the accuracy, precision, and opportuneness of the estimation of patient exposition to pollen, reducing the effectiveness of medical treatments. The advent of computer vision has brought the possibility of conducting robust pollen counts based on visual characteristics similar to those employed by palynologists and with advantage of improved speed, accuracy and reduction of variability.

Without the need of a dedicated specialist in all counting positions, the number of stations could be multiplied with the only restriction dictated from the pollen collection process, enabling the coverage of areas where the analysis was not possible before. Moreover, measurements could be taken on a daily or even hourly basis. Having the same method in all the regional stations, results would be fairly comparable.

Current vision-based methods find their foundation on the measurement of general characteristics such as shape, texture, local gray levels, or color in order to describe the pollen particle.

On the one hand, some recognition systems control the whole processing of image acquisition and pre-processing to ease the recognition task [Holt et al., 2011; Scharring et al., 2006]. However they require specialized hardware, mainly dedicated microscopes with special setup, and sometimes it involves the computation of large amount of data. The cost of additional hardware limits their application on multiple-point analysis.

On the other hand, there are systems that are completely software-based. Their advantage is that they employ the same microscope slides than in manual counting and do not require additional hardware. They are still prototypes although results are promising [Chen et al., 2006; Rodríguez-Damián et al., 2003; Boucher et al., 2002] .

Most of these methods fail in describing more specific visual pollen characteristics of the pollen, which are important on the analysis conducted by palynologists, for example, apertures. The recognition of such elements can be favorable for the discrimination of pollen that are similar in other characteristics.

1.2/ Scope of the Thesis

The current thesis investigates relevant features that enable the description of allergenic pollen taxa. The goal of the analysis is the discovery and development of features that are suitable to discriminate the taxa with recognition purposes. Special attention is paid to the recognition of apertures. By taking advantage of diverse characteristics, relevant features are expected to become more robust to natural pollen variation.

This investigation is conducted under the framework of the commercial project Personalized Pollen Profiling and Geospatial Mapping (3PGM). The project was supported by the Federal Ministry of Economics and Energy in Germany through the Central Innovation Program for Small and Medium-Sized Enterprises (Project ID KF2848901FR) from

2011 to 2013. In addition to the i3mainz institute, Bluestone Technology GmbH and health&media GmbH took part as project partners.

The aim of the project is to develop an information and analysis system for assisting the pollen allergy treatment. The main characteristic of the 3PGM project is the creation of an individual patient logbook with precise and daily information not only regarding the pollen concentration to which the patient is exposed, but also including geo-location and environmental measures. Bluestone Technology GmbH was in charge of providing the airborne pollen samples as described in Sec. 3.1. The function of health&media GmbH was the design of the user interfaces to doctors and patients for accessing their personal data. Finally the laboratory i3mainz in collaboration with the Le2i worked on the development of the pollen recognition system (matter of the present thesis) and of the framework that provides geodata analysis of the pollen concentration. An important role was the spontaneous participation of the Department of Multiphase Chemistry at the Max Planck Institute for Chemistry in Mainz, who provided the required material and facilities for the acquisition of the single-taxon dataset described in Sec. 3.2.

The current thesis focuses on the classification of the pollen taxa for the estimation of the concentration as part of the patient logbook. The characteristics of the 3PGM project set the following constraints to the image-based solution.

The approach must consider important the flexibility of the reproduction of the method at multiple regional points corresponding to multiple patients. Therefore, the employment of typical sampling slides and imaging systems, which are readily available, are necessary for the creation of the datasets. Moreover, the image processing and classification must be automated, avoiding any human intervention.

Finally, the solution must focus on the identification of the most important allergic taxa in Germany due to the regional objective of the project. Particularly, the work is focused on *Alnus, Betula, Corylus, Artemisia*, and *Poaceae*. Nevertheless, the allergenic importance of these taxa is maintained worldwide. Besides, the same procedure can be performed in the analysis of additional taxa like the pollen studied in paleontology or allergenic pollen that is relevant in other regions. For example, in Mediterranean regions, allergenic genus like *Olea*, *Cupressus*, and *Parietaria* are common. The last one can be found also in the United Kingdom. Moreover, *Ambrosia* is an allergically important pollen in the United States.

An important consideration is the taxonomical level at which the recognition is relevant and possible. The commonly accepted taxonomic hierarchy for living beings is given in Fig. 1.1. Major ranks are defined from Domain, the most general, to Species, the most specific. Intermediate sub-ranks can be also considered. For example, between Family and Genus, Subfamily and Tribe sub-ranks are commonly described.

At genus level, there are important differences in morphology and wall ornamentation among pollen types. These differences are yet advantageous for pollen discrimination in typical microscope magnifications. However, below this level, species of the same genus become too similar with only subtle differences (a property called stenopalynous in palynology), and are more difficult to discriminate in practical counts [Leopold et al., 2012; Perveen, 2006]. This similarity limits the depth of taxonomical analysis. In this case, the analysis would require more sophisticated microscopes (i.e., greater resolution and magnification) and more specialized labor capable to recognize subtle differences [Sivaguru et al., 2012]. Moreover, analysis of allergens takes place regularly at genus level, which makes unnecessary further taxonomic decomposition [Negrini, 1992].

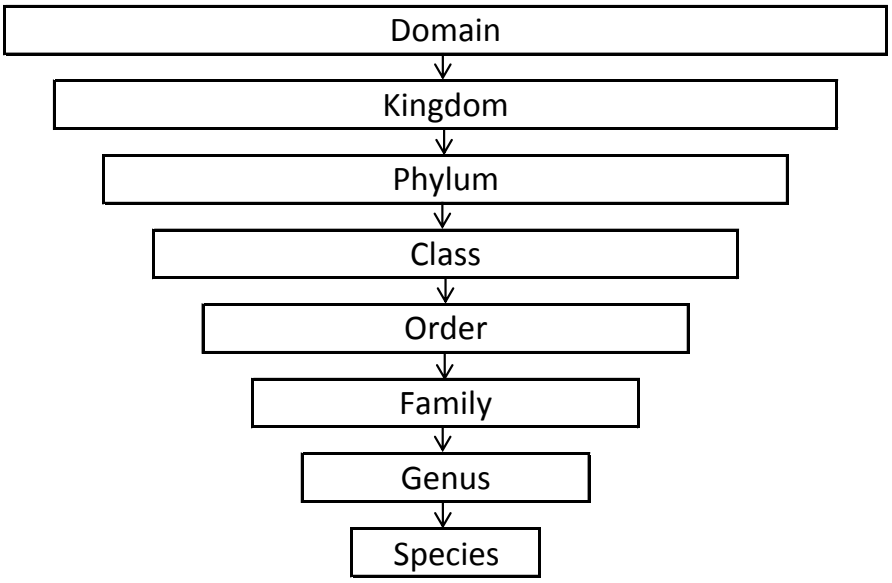


Figure 1.1: Taxonomic hierarchy of the biological classification.

In accordance with these two considerations, the present study also focus on the recognition of taxa at genus level. An important exception is the family of grasses (*Poaceae*, known also as *Gramineae*) which consist of about 750 genera (plural of genus). Most of the allergenic relevant genera belong to the subfamily *Pooideae* (around ten genera) due to its abundant airborne pollen production.

Moreover, the stenopalynous genera of the *Poaceae* present a considerable similarity making difficult its discrimination even at this taxonomic level. However, between the 80% and 95% of the cases of allergy to grasses are associated mainly to only two groups of allergens who are present in the subfamily *Pooideae*. For these reasons, counts of *Poaceae* are commonly conducted at the Family rank which does not affect drastically the efficiency of the analysis.

1.3/ CONTRIBUTIONS

This thesis proposes a recognition approach for allergenic pollen based on its visual characteristics. The contributions of the present work are:

- Development of a novel set of features that describes pollen apertures, as particular distinctive regions of the particle.
- The development of the overall method for the creation and evaluation of the detector of apertures, involving selection of important parameters.
- Proposition of a set of features that is able to describe robustly the visual characteristics of the pollen (apertures, shape, texture, and size), with particular focus on allergenic pollen taxa.

- Exhaustive quantitative and qualitative analysis of the relevance of the proposed features in the context of recognition of taxa at mainly genera level.
- Proposition of a global pollen classification method, incorporating pollen localization and segmentation. With the employment of the robust set of proposed features from different characteristics, the method is able to classify the studied taxa with an accuracy rate of more than 0.95. The developed method is planned to be used in a the 3PGM project for the automatic recognition of pollen in Germany.

1.4/ THESIS OVERVIEW

The rest of the thesis is structured as follows.

In Chapter 2, the description of the current manual process of pollen counting followed by the state of the art of vision-assisted methods for the recognition of pollen taxa are given.

Chapter 3 provides the necessary background to the characteristics of the studied pollen and describes the employed datasets.

An overview of the proposed global classification system is given in Chapter 4. This chapter also provides detailed explanation of the localization and segmentation methods. Finally, the description of the features to be analyzed is fully developed.

Chapter 5 presents the development of the aperture detector. An analysis on the selection of the correct setup, especially on the choice of the size of the region of interest, is conducted. The modular detection of apertures from different pollen types is formulated and a set of aperture features is proposed. A thorough evaluation of the detector is conducted based the count of aperture of three taxa.

The detailed analysis of the characteristic feature group is conducted in Chapter 6 through the application of feature selection methods. Based on this analysis, a subset of relevant features is proposed. At the end of the chapter, results of the global classification of pollen taxa with different combinations of the features groups are presented.

Finally the conclusions of this thesis are drawn in Chapter 7, stating the future work and perspectives.

2

State of the Art

Pollen recognition is a complex task that has been explored even before the arrival of digital computers. Palynologists, scientists specialized on pollen and spores, are responsible of the description of the pollen characteristics and learning the differences among the numerous pollen taxa. These characteristics are the foundation of the manual recognition until now. The advent of computation power brought the possibility of detecting and measuring the pollen characteristics more accurately and precisely, and analyzing multiple particles in short time.

In the first part of this chapter, a typical manual pollen recognition process is described. In the second part, a survey of the state-of-the-art methods for assisted recognition is presented. Although most of them attempt to mimic the manual process and extract some of the pollen characteristics from the 2D projection of the particle under a brightfield microscope, approached that employ alternative methods such as volumetric representation, fluorescence, spectrometry and light scattering are also introduced.

2.1/ Manual Pollen Counting

2.1.1/ Sampling and Preparation

Although there is no unique generalized sampling and counting method, there are common elements that are followed by palynologists, on which the following process is based [Mandrioli, 2000; Soldevilla et al., 2007].

First, particles are sampled from the air by means of a trap or sampler device, which is placed at open air at certain height from the base. The most common sampler is the Hirst volumetric sampler also known as Burkard trap, illustrated in Fig. 2.1. The Burkard device consists mainly of a vacuum pump and an impact drum. The vacuum pump eases the air input and control the volume of measured air. Then, the air is sucked through an intake to the impact drum. A special adhesive tape (typically based on silicon) around the drum allows the particle to adhere. A rotating mechanism attached to the drum permits to sample particles during a seven-day period at a rate of two millimeters of tape per hour. However, it is yet possible to have intermittent samples. Alternatives to the Burkard trap are the Rotorod and the Tauber samplers [Mullins and Emberlin, 1997].

Among the factors that can affect the trap performance are the presence of nearby obstacles, nearby vegetation, industrial emission and accumulation of dust.

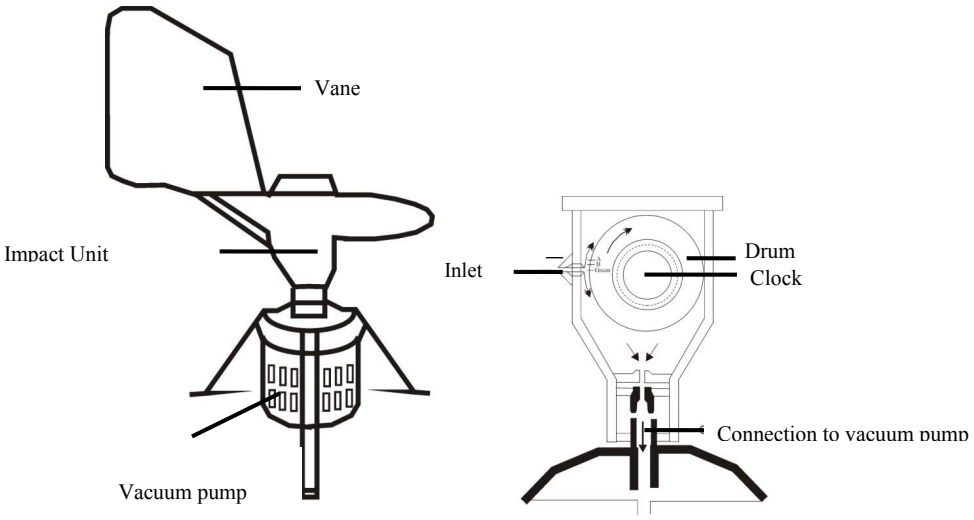


Figure 2.1: Scheme of a Hist volumetric sampler [Soldevilla et al., 2007].

The following step is the preparation of the adhesive tape for the analysis under the microscope. The tape is cut and mounted on microscope slides in sections of 48mm, corresponding to 24 h of continuous sampling. The mounting solution typically has a base of fuchsin, glycerin, and gelatin, that allows dyeing the pollen in a singular pink color. The medium is also useful to hydrate the pollen particles, which improves their morphological analysis. It is advisable to apply the mounting medium on a liquid state to reduce the number of air bubbles, phenomenon that reduces the visibility on further analysis.

2.1.2/ Slide Analysis

The prepared slide is analyzed under the microscope where the pollen identification and counting take place. Lens magnification ranges from 25X to 40X (plus and additional 10X magnification of the ocular lens). Greater magnification improves the visualization but reduces the visual field, requiring more time for the analysis of the slide.

Counting is a slow process, in which palynologist evaluates individually each particle to determine the taxon. Therefore, the whole slide cannot be analyzed completely. Pollen is identified and counted on just a portion of the sample using horizontal sweeps of length of 48mm that cover the sampled 24 h and a width depending on the microscope visual field, as depicted in Fig. 2.2a. Typically the desired portion ranges from 10% to 20%, which requires from 4 to 8 sweeps. Vertical sweeps can be employed instead if hourly information is desired as illustrated in Fig. 2.2b.

Finally, counts of each pollen taxon is expressed as the daily average concentration per cubic meter of air using the following conversion:

$$\lambda = \frac{w}{\phi_d \cdot S \cdot \eta} \cdot N, \tag{2.1}$$

where λ is the daily average concentration per cubic meter of air in $particles/m^3$, w is the width of the tape in mm, ϕ_d is mean diameter of the microscope vision field in mm, S is the

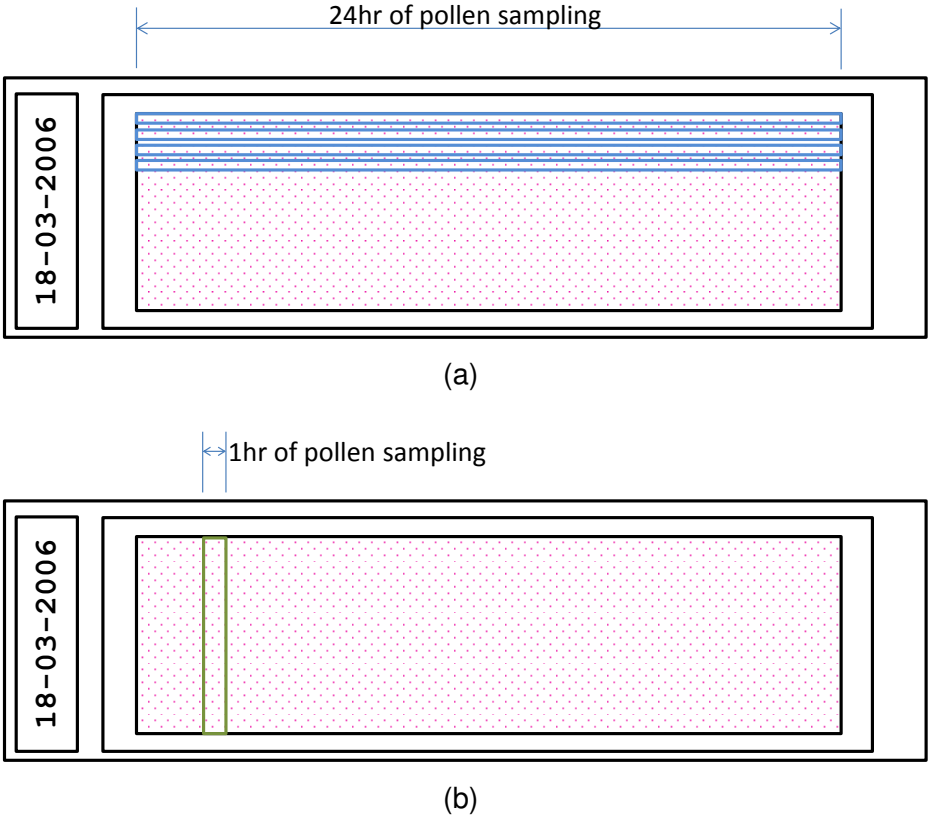


Figure 2.2: Example of daily analysis using four horizontal sweeps on a 48 mm slide for pollen counting is shown in (a). Hourly analysis employs a single vertical sweep of 2 mm in (b).

number of horizontal sweeps, η is the amount of sampled air in 24 h in m^3, and N is the number of counted particles. For the Burkard sampler $\eta = 14.4\ m^3$.

Clearly, the slowest and most difficult part is the manual taxon identification under the microscope that requires the effort and knowledge of an experienced palynologist. This process is prone to errors due to the inexperience and fatigue of the scientist. It sometimes requires the count of hundreds of particles within hours. In addition, the daily average concentration is only an estimation by extrapolation of data from just a portion of the sample. Moreover, the arduous and time-consuming labor is a cause of high analysis costs.

2.2/ ASSISTED POLLEN COUNTING

With the introduction of computer vision techniques, the possibility of new methods for pollen recognition became feasible. The capture and preservation of digital microscope images enable the subsequent analysis of the pollen, which eliminates the need of a microscope after this process is completed. Additionally, automatic and batch analysis is possible with computer-based techniques for object localization, segmentation, description, and classification.

Most of the approaches that have been conceived until now are distinguished mainly by

the pollen characteristics that are described and the type of employed feature descriptor. The following classification of the state of the art is based on the type of analyzed feature.

First, the different sources of microscope imaging are presented. Then the state-of-the-art approaches are described focused on the localization and segmentation methods, and the feature type, followed by a summary of the classification methods and results. Finally, alternative methods are briefly introduced.

2.2.1/ Microscopic Imaging

The quality and level of detail is very important for the identification of pollen. Some characteristics can be detailed only at sub-micron resolution, for example the pollen wall's exine and intine. Nevertheless, high quality of the image conveys an increase of the image acquisition time and effort, more sophisticated technology and more specialized labor. For these reasons, for a realistic continuous pollen recognition process and considering the existing technology, typical brightfield microscopes are preferred.

2.2.1.1/ Brightfield Microscope

The brightfield microscope is the most widely used and accessible in the laboratory. It is relative inexpensive and easy to use. Automated versions that can scan a section of the specimen slide into a digitized image are now popular. This function is very useful to preserve data and allows the reproducibility of the results. An example of a pollen particle image from a brightfield microscope is shown in Fig. 2.3.

Figure 2.3: Example of an alder pollen particle as seen under the brightfield microscope at a magnification of 40X and dyed with magenta.

The working principle of the brightfield microscope is the following [Wayne, 2009]: first, a light source illuminates the specimen in a parallel manner (Köhler illumination). Then, the light from the specimen passes through a sub-stage condenser and the objective lens to form the amplified image. A last additional amplification occurs at the ocular lens before the observation plane. Typical magnification of brightfield microscope ranges from 10X to 40X, plus an additional 10X given by the ocular lens. Techniques like oil immersion can increase the range from 60X to 100X.

The brightfield microscope is susceptible to diffraction patterns, caused by the destructive interference of diffracted light from the specimen to the direct light that forms the correct image on the image plane. Diffraction patterns prevent a clean and accurate image of

the specimen. Light diffraction also limits the maximum resolution of the image, which depends on the light source wavelength and the objective aperture.

Depth of field is defined as the thickness of the focal plane (perpendicular to the optical axis) in which the specimen is shown sharply. Therefore, it is not possible to observe specimens thicker than the depth of field in a single image, as in the case of pollen particles. However, it is possible to perform multiple observations at several focal planes so that the whole specimen (or the region of interest) is examined, producing an image at each plane. The result of this technique is known as multiple layers, stack of layers or optical sections. Regions that do not lie in the focal plane are blurred and fused with the on-focus image. An additional shortcoming is that the constant change of the focal plane conveys an increase on the time of the image acquisition. An example of optical sectioning of a pollen particle is depicted in Fig. 2.4.

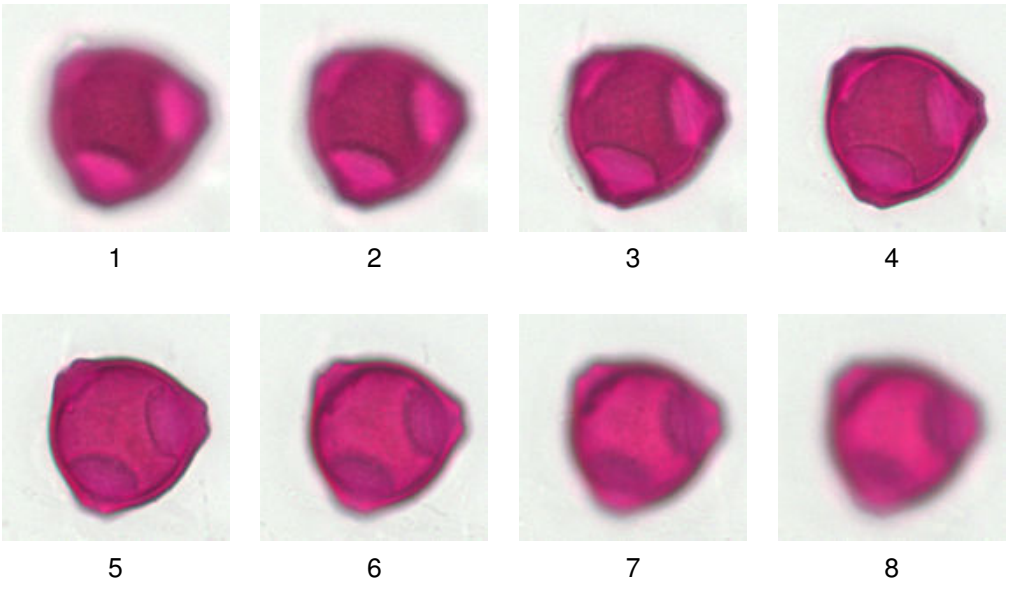


Figure 2.4: Example of optical sectioning of hazel pollen. The eight layers are extracted at different focal planes with a separation of 2 μm. Images from extreme ends of the planes become blurred. Numbers indicate the sequence of the layers.

2.2.1.2/ CONFOCAL MICROSCOPE

The confocal microscope solves the problem of optical slices of the brightfield microscope, enabling to observe exclusively a single plane without interference of blurred areas. By repeating this observation at different focal planes, it is possible to obtain a 3D dataset of the specimen with rich information of its spatial structure.

Confocal microscopes use a pinhole aperture (confocal) in a conjugated plane between the objective and the detector such that the latter receives light information from a single point of the specimen [Semwogerere and Weeks, 2005]. Additionally, a laser beam is focused only on the same observed point. An example of a pollen particle image from a confocal microscope is shown in Fig. 2.5. One of the limitations of confocal microscopes is the drastic reduction of light impacting the detector, fact that decreases the signal-to-noise ratio and increases the acquisition time.

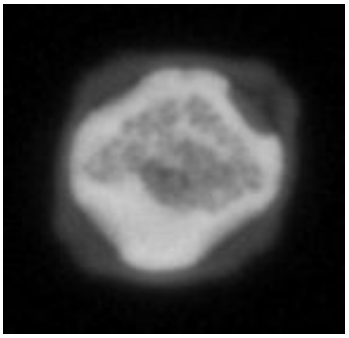


Figure 2.5: Alder pollen using confocal microscope [Ronneberger, 2007].

2.2.1.3/ FLUORESCENCE MICROSCOPE

An alternative source of information from the specimen is the fluorescence. Fluorescence is the property of some substances to absorb light and emit it eventually usually with a different wavelength. Fluorescence microscopy allows to reveal and contrast differences in the specimen that are unnoticeable in traditional microscopy and it is special useful in biological sciences.

The operation principle is to irradiate the specimen with a light in a narrowed wavelength [Lichtman and Conchello, 2005]. The very weak emitted light of the specimen is gathered by a detector. An example of a pollen particle image from a fluorescence microscope is shown in Fig. 2.6. A typical limitation of fluorescence microscopes is the very low ratio between the detected light and the source light. Additionally, the photobleaching occurs when the fluorescent property of the substances fades out after the exposition to light, preventing continuous or repeated observations. Confocal microscopy advantages are also commonly employed in combination to fluorescence microscopy.

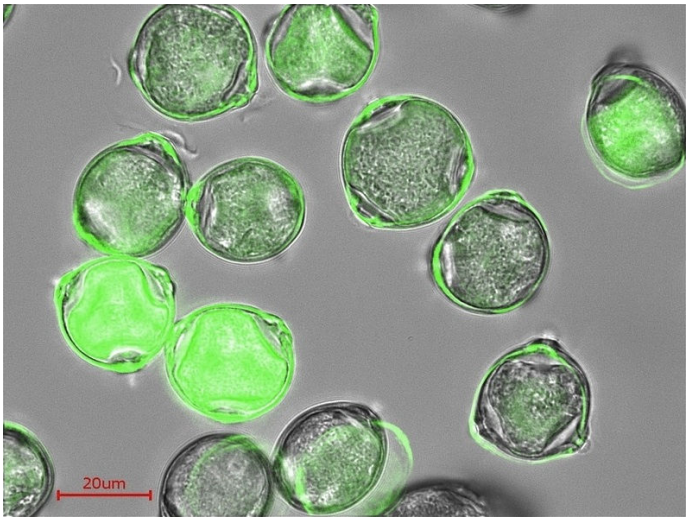


Figure 2.6: Example of birch pollen colored by a fluorescence microscopic picture. Chemical composition is the responsible of the difference in brightness[1].

[1] Picture credit © Christoher Pöhlker/MPI for Chemistry from http://www.mpg.de/1170854/.

2.2.1.4/ SCANNING ELECTRON MICROSCOPE

The principle of a Scanning Electron Microscope (SEM) is the interaction of an incident electron beam on the specimen [Egerton, 2005]. Ejected electrons are collected and transformed into an image of the specimen morphology with a resolution up to 50 nm with lens magnification up to 30,000X. Additional information of the chemical composition and the crystalline structure and orientation can be extracted by the analysis of ejected backscattered and auger electrons, heat, and X-rays. An example of a pollen particle image from a SEM is shown in Fig. 2.7.

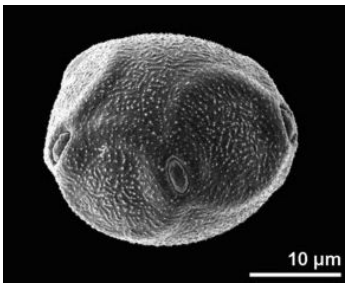


Figure 2.7: Example of an alder pollen particle as seen under the SEM microscope. Picture credit: Halbritter H., *Alnus incana* in [Buchner and Weber, 2014].

Although the level of detail is higher and the variety of information is wider than with other microscope technologies, their application to batch pollen recognition is limited because the acquisition time is in the order of minutes. Furthermore, the high cost of the device, the required infrastructure and the specialized operation is not affordable for demanding pollen analysis.

2.2.2/ SEGMENTATION AND LOCALIZATION

Real airborne slides prepared from an airborne particle sampler contain many types of objects additionally to pollen particles, such as vegetation debris, spores and insects. The degree of contamination is very variable and depends on the surroundings of the sampler. All these objects are cluttered in the slide and increase the difficulty of finding isolated pollen particles. In such cases, a localization step is required to identify the position of the pollen particle in the image and to reject the debris.

After identifying the potential pollen and knowing the location, most of the methods require the trace of the particle contour in order to determine the area of the image that belongs exclusively to the particle by means of segmentation methods. Typically, this step is necessary for the computation of features, especially those describing shape.

For color and texture identification, a sample of the particle area is sometimes employed to compute the features with the risk of adding part of the background or leaving important regions out of the computation. General segmentation techniques can be applied to this end, however, model-driven methods can be employed too.

The most useful characteristics to identify pollen from the debris are the near-circular shape (especially on allergenic taxa) and the pink color due to the dyeing. The dyeing affects almost exclusively pollen, leaving most of the debris and spores in their original color.

Rodríguez Damián *et al.* applied a method based on the Hough transform and active contours to extract the nearly circular shapes from airborne samples [Rodríguez-Damián et al., 2003]. The slide image was pre-processed as follows: first, noise was reduced with a mean filter, and then a Sobel operator was applied for edge detection. The binary borders were obtained by thresholding with Otsu's method. Imperfections were eliminated with the erosion operation. Finally, the result was subtracted from the original binary image and borders were refined with a thinning algorithm.

Pollen particles were recognized applying Hough-transform-based circle detection. The expected range of radii of the pollen was employed to restrict the circle search. The sensitivity to imperfect circles was controlled by a threshold on the parameters. The segmentation of the pollen employed an active contour method, namely snakes. They started from the Sobel-filtered binary image and used the detected circles as first contour estimations. There was no quantitative evaluation of these preprocessing since segmented pollen was employed in a global classification process.

Landsmeer *et al.* based the detection on color and circular similarity [Landsmeer et al., 2009]. They employed a low resolution version of a registered image from a stack of 24 layers of airborne pollen slides. First, they assigned a value to pixels regarding their similarity to the pollen color (color similarity transform) employing a quadratic discriminant classifier in a hue-saturation color space, and eliminating the illumination variance. Then, they detected circles on the transformed image using also the Hough transform similarly to Rodríguez Damián *et al*. They also used the Hough transform values to determine the sharpest layer of the stack. Considering the cluttered contamination of the slides, recall of 86% and precision of 61% seem a good result.

The French-Spanish ASTHMA project combined localization and segmentation [Boucher et al., 2002]. A search of a sharp layer from a stack is performed. First, colors on the image are coded using the best color axes after PCA analysis. The segmentation employed a split-and-merge technique with Markovian relaxation in combination with chrominance and luminance properties. They achieved a recall of over 90%. There is no information about the content of the debris on the slide to evaluate the performance correctly. Additionally, the use of luminance could make the technique not robust to changes on the scanning process.

Ranzato *et al.* applied a difference-of-Gaussians filter for the localization of pollen from cluttered slides [Ranzato et al., 2007]. The filter was applied to sub-sampled versions of the image to extract eight interest points on the border, related to circular shapes. The final bounding box containing the particle was obtained from a circle fitted to the selected points.

The set was reduced based on size and brightness criteria. The method achieved a recall of 93.9% and precision of 8.6%. They disregarded this low precision because of the excessive amount of debris. Moreover, they proposed to reject the intrinsic high number of false alarms on the classification stage.

Ronneberger *et al.* employed their own voxel-wise vector based gray-scale invariants for localization of pollen particles, which is based on the detection of circles [Ronneberger et al., 2007]. Energy minimization methods were selected for the segmentation of particles. For brightfield microscope images, snake segmentation was applied to weighted-edge images of the particles. For the case of 3D surfaces of confocal images, they scaled down the data and applied the graph cut algorithm. There were no individual quantitative results of this stage.

Allen *et al.* reported also the use of Sobel edge detection in combination with dilation, erosion and filling for localization of blobs, which represented potential pollen particles, and their segmentation [Allen et al., 2006]. Blobs of size different from pollen were rejected.

Chen *et al.* employed a two-step thresholding for segmentation [Chen et al., 2006]. First, using the automatic triangle threshold, a coarse approximation of the particle was extracted. Then, a fine result was obtained with the application of the automatic IsoData threshold. Finally, gaps were fixed using hole filling.

2.2.3/ Shape Features

The recognition of the pollen shape has been widely used due to the diversity of available geometric and morphological features already proven in object recognition. The ASTHMA project, for example, employed the following global shape features: area, perimeter, compactness, equivalent circular diameter, convex hull area, convex hull perimeter, concavity, convexity, solidity, moments of inertia and eccentricity [Boucher et al., 2002]. The computation of several of these features is based on the perimeter and area of the shape, which suggests redundancy in addition to a direct association with the size of the particle.

Rodríguez Damián *et al.* computed global shape features on the extracted pollen contour: area, perimeter, roundness, centroid, mean, maximum and minimum distances to the centroid, ratios between these distances, diameter, number of holes, boundary roughness and radius dispersion [Rodríguez-Damián et al., 2003].

A very similar set was created by Chen *et al.* It consisted of area, perimeter, circularity, mean, maximum and minimum distances to the centroid, ratios between these distances, diameter, radius dispersion, length after skeletonization, and seven binary Hu moments [Chen et al., 2006].

Zhang *et al.* computed seven parameters from central moments, which are invariant to translation, rotation and scaling [Zhang et al., 2004].

France *et al.* employed the Paradise Neural Network to describe the shape of the pollen [France et al., 1997]. This method was originally designed for the recognition of hand gestures and later generalized to the classification of other objects. The architecture of this kind of neural network comprehends three layers. The first layer is in charge of detecting horizontal and vertical lines at two frequencies. The second layer automatically builds templates based on patterns of the extracted features that represent the pollen particle. Finally, the patterns are classified into automatically created classes in the third layer.

These global measures summarize the shape information in few statistics. However, a measure that describes more detailed the pollen outline has not been employed yet. This type of analysis could, for example, consider characteristic shape patterns of different pollen types.

2.2.4/ Texture Features

The only group that has deeply studied texture features for pollen recognition is the Massey University in its Classifynder/AutoStage project, which is oriented to paleontological studies. Li *et al.* computed five Haralick measures, which are well known texture

descriptors, from the Grey Level Co-occurrence Matrix. The measures were angular second moment, contrast, variance, inverse difference moment and entropy [Li et al., 2004].

Also from the Massey University group, Zhang *et al.* proposed the application of the even-symmetric Gabor filter to the image at six orientations and nine scales [Zhang et al., 2004]. Rotational invariant features were computed as the average response of the filters corresponding to different directions. The final employed descriptors were the average mean and standard deviation of the invariant features.

Chen *et al.* normalized the gray-level pollen images to the mean of the background [Chen et al., 2006]. Then, they applied a set of statistical gray-level features on a central patch on the image. They proposed a feature computed as the mean square difference between the histogram of the evaluated image and the mean image of each of the studied pollen types. Additionally, they computed the mean and standard deviation of the gray values and of the gradient intensity, the standard deviation of the Laplace image, the minimum gray value of the cell wall, and seven gray Hu moments.

2.2.5/ Local Features

Features that describe locally the interaction between gray levels of the image have been brought from object recognition. Ranzato *et al.* applied a set of kernels from Koenderink and van Doorn to a scale pyramid of the image as a base for shift and rotation invariant local jets [Ranzato et al., 2007; Koenderink and van Doorn, 1987]. Their approach did not required previous segmentation of the particle.

Ronneberger *et al.* designed local 3D gray level descriptors invariant to translation, rotation and global deformations [Ronneberger et al., 2007]. The descriptor is founded on the Haar-integration over transformation groups and on the use of parametrized kernel functions.

Local features seem a powerful method of describing complex morphologies. However, the need of volumetric data limits the feasibility of implementation. It would be interesting to evaluate strategies based on local features in combination with other types of features, like shape and texture.

2.2.6/ Color Features

The color of pollen seems to be an important factor to discriminate pollen from other particles and debris, due to the characteristic pink color after being dyed. However, a lack of a standard process and variability on the dyeing substances hinder its use on the recognition of taxa.

Color representation of pollen taxa could be useful only under the process conditions on which the color information was modeled. Furthermore, there are no solid results yet that support a recognizable color difference among taxa. For this reason, color is mainly employed only for localization of pollen and for rejecting debris. Until now, only the ASTHMA project has added the mean red, green and blue colors of the particle to the set of features [Boucher et al., 2002].

2.2.7/ Apertures and Other Particular Characteristics

Particular characteristics, which are distinctive of pollen, have been also employed to differentiate classes. For example, palynologists are able to recognize aperture, perforation, foveola, stria, granulum, reticulum, etc. Sometimes, they can be a key factor in discriminating similar genera. One of the most important characteristics is the aperture because of its visibility and discrimination power. The type and amount of apertures are particular of the pollen taxon. Despite its importance, few works exist on detection of apertures.

Apertures are morphological distinctive regions located on pollen external wall that undertake mainly the function of germination. Apertures typically show thickening of the pollen wall around their perforation. They can be discriminated morphologically from the rest of ornaments in the pollen wall due to the bigger size. A detailed classification of the aperture types can be found in Sec. 3.3.1.

The ASTHMA project developed several algorithms to detect some specific characteristics of the pollen such as apertures, and cytoplasm [Boucher et al., 2002]. However, each algorithm was designed for a specific characteristic, taxon and viewing angle. The requirement of multiple algorithms is impractical for the analysis of multiple taxa.

They employed a combination of methods to detect and locate the characteristic such as thresholding, Laplacian of Gaussian, and region segmentation. Then, each type of appearance of the characteristics was described by the color, size and shape. Finally, a validation on multiple layers was conducted.

Chen *et al.* followed a similar approach [Chen et al., 2006]. They developed algorithms for different appearances of aperture of birch and mugwort pollen implying the same inflexible need of designing algorithms for new types.

The Hough transform was applied on the pollen image for the detection of circular apertures (pores). The analysis was adapted to a specific range of radii. The maximum of the resulting image was considered as a feature itself. Then, the center of the aperture was determined by thresholding this image. The final aperture count from three different thresholds, oriented to different performance criteria, was employed as a feature.

They analyzed birch apertures on the particle image border using the polar transform of the image. A template matching strategy was applied on the transformed border in order to find the apertures. The similarity measure was the cross-correlation, and thresholding was used to find the aperture location. Similar to the circular aperture detection, the final features were the three different counts obtained by the variation of the threshold in addition to the maximum of the cross-correlation image.

For the detection of mugwort apertures on the borders, Chen *et al.* used also the polar-transformed image. First, the transformed image was aligned to the particle wall. Then, a gray level intensity profile was extracted from the particle wall. Apertures were obtained by thresholding the profile and regarding the high peaks as positives. The total count of apertures was considered as a feature. An additional feature was computed as the highest peak of the frequency spectrum of the intensity profile.

Aperture features were fed to a global pollen classification system with recognition at taxon level. Individual evaluation of the detectors is not shown, limiting the performance comparison. However, it is evident that each set of features is very dependent on the aperture type and the employed method. A solution that employs a unified aperture descriptor would be more practical and powerful. The use of a descriptor that can describe

the spatial pixel arrangement of the different aperture patterns and shapes seems more discriminative.

Focused on only one pollen type, France *et al.* developed an algorithm for the detection of the apparent double edge of hawthorn pollen [France et al., 1997]. They transformed the pollen image into polar-coordinate image and found the edges using a Gabor detector. After blurring the result, they applied a filter designed specifically to respond to the double edge. Treating the inverse of the resulting image as an energy landscape, they applied an active contour detector (snakes) using the landscape slope to determine the image forces. This operation yields to a percentage value of the double edge. Finally, a threshold is employed to determine the necessary percentage value for regarding a detection. This method is too specific to one type of pollen and does not seem to be replicable to typical allergenic pollen.

The exploration of methods for aperture detection appear important and little work has been done. Focus on the morphological variations due to the pollen type must be considered if a unified method is to be designed: for example, the common patterns present in the types of aperture but still different from the rest of the wall. This requires an analysis of the aperture in small visual components.

2.2.8/ Classification Methods and Results

The dataset of the ASTHMA project consisted of 30 pollen taxa [Boucher et al., 2002] and 350 pollen grains for an average of 11.6 particles/taxon. Interestingly, they employed a knowledge-based classification instead of typical statistical methods. At the first step, the shape and color features were employed to reduce the list of possible taxa. At the second step, specific characteristics were detected based on the reduced list in order to recognize specifically the pollen taxon. The method allowed the analysis on multiple layer in order to validate the detections. The final accuracy of the classification was 77%. However, the average of 11.6 images/taxon in the dataset looks small for capturing all the variations of the taxa.

Rodríguez Damián *et al.* tested their system on three pollen taxa and 185 particles [Rodríguez-Damián et al., 2003]. The classification method was a minimum-distance classifier and it was evaluated within a leave-one-out scheme. Best features were selected using a modified floating search. Results showed a classification rate of 85.62% with five selected features. The performance is good considering only shape features and a moderate dataset size.

Li *et al.* employed the Fisher's linear discriminant method to classify four pollen taxa [Li et al., 2004]. Although the performance was 100% of accuracy, the test data appeared to have very distinctive texture patterns and a small inter-class similarity, reducing the difficulty of the task. Results leave unknown the performance on typical allergenic pollen.

Zhang *et al.* employed a multi-layer perceptron (neural network) with one hidden layer [Zhang et al., 2004]. A set of 36 images from five pollen types was evaluated including three artificially rotated versions. The dataset was split into two parts for training and testing. They correctly classified 97.7% of the testing samples. However, the evaluation of a more significant amount of samples and the consideration of important allergenic taxa would have been desirable.

Allen tried to reduce the features from Li *et al.* and Zhang *et al.* in addition to wavelet

features using Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) [Allen, 2006]. However, results showed no improvement by reducing the feature space dimensionality.

Chen *et al.* used a dataset of 254 images from birch, mugwort and grass pollen [Chen et al., 2006]. A forward feature selection yielded a reduction from 47 to 12 features. The best set consisted of seven general shape features, four aperture features, and a statistical gray-level feature. They achieved a 97.2% of accuracy with a Liner Normal Classifier (LNC) within a five-fold cross-validation scheme.

They stated that mainly statistics related to distances of the border to the centroid (size dependent), circularity and length after skeletonization were among the selected features together with the mean of gradient intensity. They indicated also a classification error reduction of 2.3 percentage points when some pore features were included.

The Paradise Neural Network employed by France *et al.* performed the description and classification of the pollen without the need of additional data or methods [France et al., 1997]. The drawback is the automatic creation of the output classes according to the analysis, similar to a clustering approach. They employed 1800 images of hawthorn pollen split into training and testing datasets. Because the dataset included mostly debris, more than the half of the created classes contained solely debris.

The system discriminated correctly 79.1% of the pollen from the debris. However, 72% of the debris could not be classified due its high variability. Results suggest that the method may be insufficient for taxonomic classification. For the recognition of the double edge on the same datasets, France *et al.* achieved a recall of 70% of recognition. Unfortunately, the method was not tested on additional pollen types.

Ranzato *et al.* reached the best results limiting the experiment to three pollen types: ash, alder and pine [Ranzato et al., 2007]. The Bayesian classifier was used to model a mixture of Gaussians. The feature space was reduced by means of Fisher's linear discriminants. For the validation of results, 100 experiments were conducted using randomly 10% of the dataset for testing. The global accuracy was 82.6% including the recognition of non-pollen particles. Accuracy reduced to 64.9% when eight pollen taxa were classified.

Ronneberger *et al.* studied 33 pollen taxa [Ronneberger et al., 2007]. Employing a SVM classifier and the Radial Basis Function (RBF) within a leave-one-out evaluation scheme, they achieved a precision of 98.5% and recall of 86.5% on dataset of 22700 particles with multiple layers after rejecting results from uncertain particles. This is clearly the best result considering the amount of taxa and sample particles. Nevertheless, the requirement of multiple-layer image for the computation of the invariant seems hinder its use on an agile system.

The presented classification systems tend to give stronger importance to one or two of the features types. It would be important to evaluate the discriminative relevance of the features and to test the joint performance of the combination of different sets of features.

2.2.9/ Alternative Methods

Alternative methods for pollen identification benefit from their particular luminous properties such as autofluorescence, multispectral profile and scattering reflection.

Mitsumoto *et al.* extracted the Blue to Red ratio (B/R) of nine pollen species applying

two different methods of measurement [Mitsumoto et al., 2009]. In the first method, they placed sequentially blue and red color filters in front of a digital camera. In the second method they computed the whole color spectrum in the range of 400 to 700 nm by means of a band-pass filter and averaged blue and red subranges. Additionally, they estimated the size of the particle using two methods, a sheath-flow cell system and a particle size analyzer. Comparable results were obtained by both methods. A B/R-size plot showed that the nine species were partially identified according to these two variables. However, there is overlapping on the taxon clusters preventing a perfect classification. Although there is no quantitative results, results suggest that autofluorescence can be an important discriminant of pollen taxa.

Dell'Anna *et al.* experimented with Fourier transform infrared spectrometry for imageless recognition of 11 pollen taxa [Dell'Anna et al., 2009]. This technique has been previously employed for the characterization of microorganisms. They gathered eight mid-infrared spectra of two alternative modes: from a single pollen particle (five spectra) and groups of multiple particles of the same taxon (three spectra). Results showed that the taxa could be grouped into different clusters with a hierarchical clustering using the multiple-particle dataset.

On a second experiment using a K-nearest neighbor classifier within a leaving-one-out cross-validation scheme, they achieved the best result using the single-particle dataset for training and testing with an accuracy of 84% for 55 spectra, having for some taxa precision below 80%. Additionally, they found that the most discriminant range of the spectrum was between 850 to 1800 cm^{-1}. Although it was shown that part of the spectra is discriminant among some taxa, the training set seems too small to generalize a classification model.

Kawashima *et al.* designed an automatic complete system to sample particles from the air, to filter most of the pollen, and to recognize three pollen taxa [Kawashima et al., 2007]. The principle of the optical section is the detection of the forward and sideways scattering of the pollen particles impacted by a ribbon laser beam. They stated that the intensity of the scattered light was influenced mainly by the reflectance and size of the particle, although shape and surface roughness was also a factor. A characteristic region in the forward and sideways space was determined for each of the pollen types.

Identification results for ten-days sampling were close to manual counts by specialists. In this experiment, a more precise method to determine the region limits would have been desirable. Further investigation is needed to determine the factors affecting the differences of the scattering.

The specialized equipment in these alternative methods limits their application to agile pollen recognition.

3

Pollen and Datasets

Pollen databases are important in the development of a taxa recognition system. The employed methods and features depend strongly on the type of database. The first subject to address is the source of the samples, whether the pollen is sampled from a real environment or prepared in the laboratory. The second subject is the imaging process, which implies the type of microscope, the mode (brightfield or fluorescence), magnification, etc.

In the present study, we analyze the five most important allergenic pollen taxa in Germany. The criteria to determine their importance were the allergenic potency and the predominance in the environment. The allergenic potency of a pollen taxon depends on the type of allergens that the particle carries, which are the direct responsible of allergic reactions on the human being. The predominance or frequency of the pollen is estimated by the annual counts from multiple stations in a specific region. In Germany, the authority in this subject is the German Pollen Information Service Foundation (*Stiftung Deutscher PollenInformationsDienst* also known only as PID), who works with about 45 measuring stations across the country. The frequency data, which have been gathered by the PID for more than ten years, in combination with the analysis of the allergenic potency have yielded to the list of the most important allergenic pollen. In this study, five out of the seven taxa from the list have been considered, amounting the 99% of share, are alder (*Betulaceae Alnus*), birch (*Betulaceae Betula*), hazel (*Betulaceae Corylus*), mugwort (*Asteraceae Artemisia*), and grass (*Poaceae* or *Gramineae*). The share for each taxon [Bergmann, 2014] and their allergenic potency [Réseau National de Surveillance Aerobiologique, 2014] is shown in Table 3.1.

In the rest of this chapter, the two datasets that were employed for the design of methods,

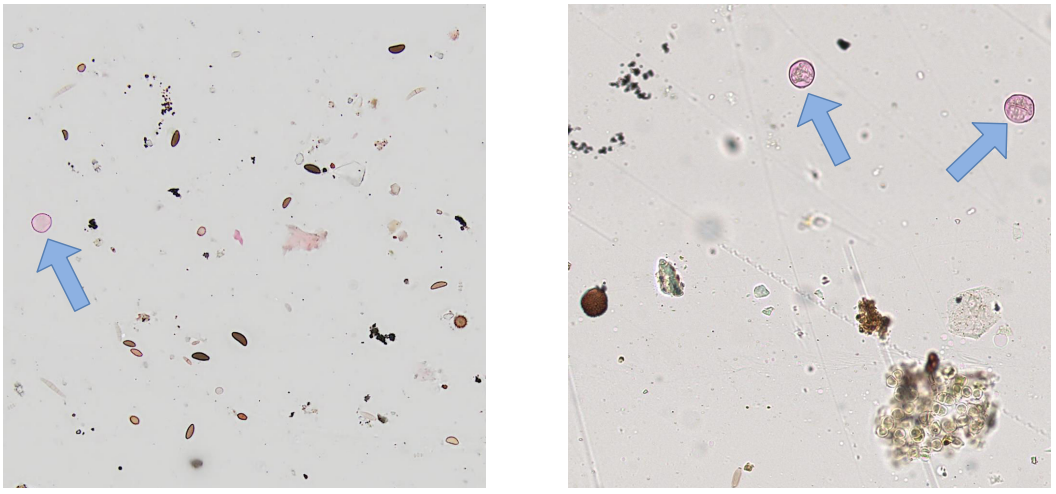
Table 3.1: The five studied allergenic pollen taxa. Shares of the last two years amount for about 99%. There is not a universal consensus on how to express the allergenic potency. Therefore, only rough levels can be safely described in the following scale: 0-no potency to 5-very high potency.

Pollen taxon	Share 2013[%]	Share 2012[%]	Allergenic potency	Main allergen
Alder	26.0	14.4	4	Aln g1
Birch	51.0	59.1	5	Bet v1
Hazel	5.5	3.6	3	Cor a1
Mugwort	1.0	2.3	3	Art v1-3
Grass	15.9	19.9	5	Lol p1-11

analysis of features and experiments are described. Then, the most relevant palynological characteristics of the studied taxa are explored.

3.1/ AIRBORNE DATASET

This dataset represents the type of image that is expected from a real airborne sampling, for example, using a Burkard trap (*cf.* Sec. 2.1.1). It is cluttered with debris, outnumbering the pollen particles. The spatial distrubion and the angular position of the particles are completely random and is not controlled. Two different slides were dyed with fuchsin. One slide was scanned by an Olympus VS120 brightfield microscope with an objective magnification of 20X and an image resolution of 0.35 μm/pixel. The second was scanned with a Leica SCN400 brightfield microscope with an objective magnification of 40X and an image resolution of 0.25 μm/pixel. An extract of both slides is shown in Fig. 3.1 as an example. As expected, a quick inspection reveals that the rate of potential pollen particle (pink circles) to debris (brown or irregular shapes) is too low.



(a) Olympus microscope (b) Leica microscope

Figure 3.1: Extracts of slides from the airborne dataset: (a) using the Olympus VS120 microscope and covering a square area of side length 350μm and (b) using the Leica SCN400 microscope and covering a square area of side length 250μm. Potential pollen is pointed out by a blue arrow.

The resulting two images of this dataset is employed exclusively for the development of localization and segmentation algorithms in real data. The number of pollen, estimated in around 39 particles in both images, and the taxon variety is too low for the validation of taxa classification in these slides. The estimated total number of sampled particles, including debris is 500.

3.2/ SINGLE-TAXON DATASET

The second dataset consists of slides containing pollen from a single type among the five considered ones. Pollen was previously gathered and classified by palynologists. Later in the laboratory, pollen is spread on the slide, dyed with fuchsin and rehydrated. This type of preparation avoids the contamination from different particles, especially debris. The spacial distribution and the angular position of the particles is also random as in the case of the airborne dataset. Nevertheless, the number of particles and the dispersion method can be controlled in order to prevent cluttered slides. The angular position cannot be fixed. Moreover, the distribution of the particles on the slide does not indicate the sampling time slot as in the case of the Burkard sampler.

The slides were scanned using a brightfield microscope Keyence BZ-9000 with magnification 40X and a resolution of 0.26 μm/pixel. Particles were stained with fuchsin dye. They were manually located and individual image snippets were cropped. The following amount of particles for each taxon were extracted: alder 115, birch 136, hazel 52, mugwort 120 and grass 132.

Pollen particles of hazel stuck each other excessively, creating clutter pattern impossible to split into individual snippets. This problem reduced their availability. This dataset was employed for the analysis of each taxon, especially the application of different features and the classification experiments. An example of one of the original slides and extracted snippets is shown in Fig. 3.2 and 3.3 respectively.

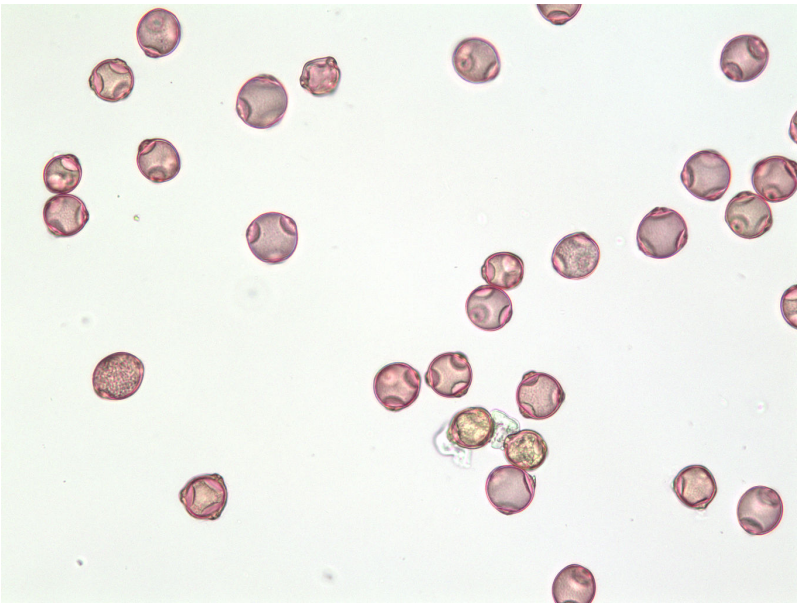


Figure 3.2: Extract from a slide of the Birch single-taxon dataset. All the particles are known to be pollen.

3.3/ PALYNOLOGICAL INFORMATION

In palynology, there is no definite consensus yet on the description of pollen and spores. Therefore, the terminology is still not exact and there are still controversial terms. How-

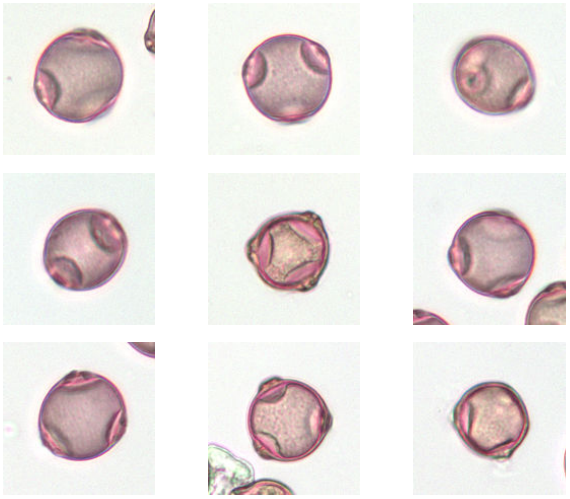


Figure 3.3: Example of snippets extracted from image in Fig. 3.2.

ever, the image-based analysis of features and recognition of taxa is not affected by this disagreement. In the current section, the most accepted terminology for pollen description is described.

3.3.1/ Morphological Information

Morphological characteristics have been until now the main criteria for the recognition of pollen taxa. Since manual analysis is carried out on the amplified image of a microscope, pollen description is mainly based on characteristics that are observable by this means. Moreover, the introduction of more powerful imaging systems, such as SEM, has allowed the recognition of tiny characteristics below the order of μm.

Shape. If the particle is seen as a spheroid, the polar axis and the equatorial plane can be identified. The volumetric shape of the pollen can be expressed as a ratio between the Polar axis length (P) and the Equatorial diameter (E) [Erdtman, 1943]. Fig. 3.4 shows the topology of the polar axis and the Equatorial diameter in a pollen particle. Determination of the spheroid poles (polarity) depends on the patterns on the morphology formed by the union of multiple particles, called dispersal units, during the formation period. Therefore, the determination of the polarity is not straightforward and depends on the particular taxa. Usually, key information of the position of characteristic of the particle, like apertures or furrows, is needed beforehand.

Shape categories according to the P/E ratio are given in Table 3.2. For the case of the spheroidal shape, it is difficult to find a perfect P/E ratio = 1, therefore some variance is allowed. If not enough information about the morphology of the particles and the position of key regions is known, it is not possible to identify between categories with reciprocal P/E ratio (Perprolate - Peroblate, Prolate-Oblate, etc.).

Additionally, it is common to describe the outline of the particle as seen perpendicularly to the equator, in direction of the polar axis (polar view). Although many pollen types are seen as circular when are completely hydrated, it is convenient to describe also the outline when the particle is partially or entirely dried, which are the most frequent states found during the analysis. Most common outline shapes are circular, polygonal, triangular,

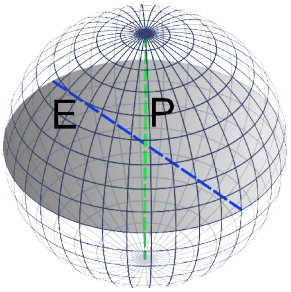


Figure 3.4: Topology of the polar axis P (green dashed line) and the Equatorial diameter E (blue dashed line) in a pollen particle. Their definition is based on the determination of the Equatorial plane (gray surface)[2].

Table 3.2: Shape categories based on the P/E ratio.

Shape category	P/E ratio range
Perprolate	>2.00
Prolate	1.33 - 2.00
Subprolate	1.14 - 1.32
Prolate-spheroidal	1.01 - 1.13
Spheroidal	1.00
Oblate-spheroidal	0.88 - 0.99
Suboblate	0.75 - 0.87
Oblate	0.50 - 0.74
Peroblate	< 0.5

lobate, or irregular.

Size. Size of most of the pollen ranges from 10 μm to 100 μm. Pollen taxa outside this range are exceptional. The pollen size varies from taxon to taxon. The lower the taxonomical level is analyzed, the smaller size variance is found. There is also a natural variation among individuals of the same taxon. Therefore, the size range of a particular taxon can be determined statistically. Ranges are given typical at family or genus level, although palynologists do not always agree. Differences could be due to the fact that not all sub-taxa are represented in the sampling. A reason may be the regional availability of the taxa.

Ornamentation of the pollen wall. This characteristic describes the morphology of the external wall of the pollen. The pollen wall is formed by two main layers, the exine and the intine. A detailed stratification of pollen wall is depicted in Fig. 3.5. When the ornamentation takes place at the exine, it is called sculpturing. Different structures and sub-characteristics on the wall yield to wide variety of patterns. Some of them are so small that they are hardly visible at magnifications smaller than 40X.

Most common ornamentation types are described below and shown in Fig 3.6. The reminding types can be found in the extensive glossary of pollen and spore terminology in [Punt et al., 2007], from where the following definitions where taken.

[2]Image modified from Geek3, *sphere wireframe* in commons.wikimedia.org under CC BY-SA 3.0.

- Granulate. Composition of rounded small elements (granules). When the elements are smaller than 1 μm in all directions, the type is called scabrate.
- Rugulate. Composition of elongated elements larger than 1 μm forming an irregular pattern.
- Echinate or microechinate. Composition of long tapering elements (spines) larger than 1 μm.
- Reticulate. Composition of spaces (lumina) wider than 1 μm and delimited by border elements (muri), forming a pattern similar to a network.
- Foveolate. Composition of rounded depressions with diameter greater than 1 μm. Foveolae are separated by distances greater than their breadth.
- Psilate. Having a smooth surface.

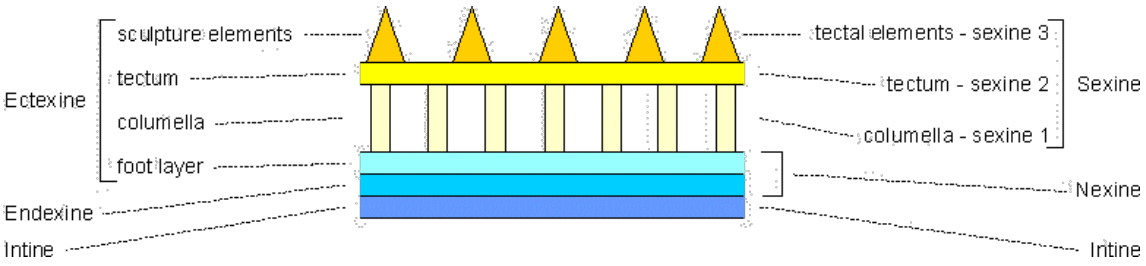


Figure 3.5: Stratification of the pollen wall. Differences on the names between the right and the left correspond to two different naming systems. Image from [Punt et al., 2007].

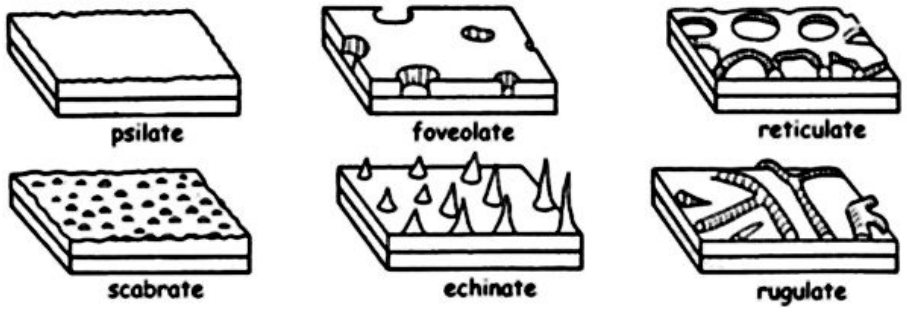
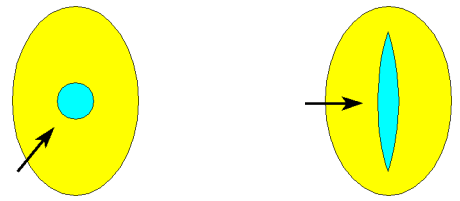


Figure 3.6: Common ornamentation types of the pollen wall. Image (shortened) from [Institute of Plant Sciences, University of Bern, 2003].

Apertures. Apertures are morphological distinctive regions located on pollen external wall typically thinner than the sporoderm. Since apertures fulfill the specific function of germination (sometimes also for the transference of substances) and are bigger than the ornamentation elements (> 1 μm), they are classified independently to the ornamentation types. The type and amount of apertures are an important factor for the discrimination of taxa, and sometimes can be the determinant for distinguishing between two similar taxa. Together to their accompanying elements, apertures can be often observed with a common brightfield microscope due to their size.

According to their morphology, apertures can be classified as (cf. Fig. 3.7):

- Pore. Circular or elliptical shape. Length/breadth ratio less than two. Ulcus is a special pore defined as an ectoaperture at the distal or proximal pole (refer to definitions below).
- Colpus. Elongated shape. Length/breadth ratio greater than two.

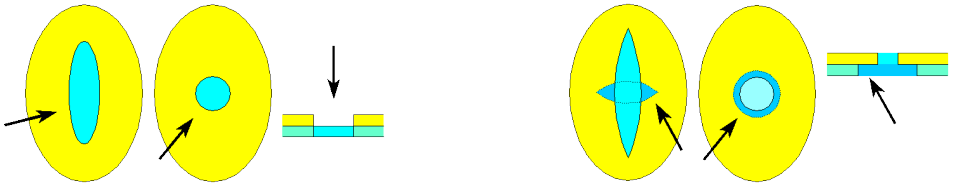


(a) Pore b) Colpus

Figure 3.7: Type of apertures according to their morphology. Images from [Punt et al., 2007].

According to their position on the pollen wall, apertures can be (cf. Fig. 3.8):

- Ectoaperture. Located on the outer layer (sexine / ectexine) of the sporoderm. For example, ectocolpus and ectopore.
- Endoaperture. Located on the inner layer (nexine / endexine) of the sporoderm. For example, endocolpus and endopore.



(a) Ectoaperture (b)Endoaperture

Figure 3.8: Type of apertures according to the layer position on the pollen wall. Images from [Punt et al., 2007].

Apertures can also consist of combined morphologies at different layers of the wall. The most common classes are (cf. Fig. 3.9):

- Colporate. Combination of an ectocolpus and one or more endoaperture.
- Pororate. Combination of an ectopore and an endopore in a not congruent way.
- Colpororate. Combination of an ectoaperture, a shorter lolongate mesoaperture (longitudinally elongated middle area) and a lalongate endocolpus (transversely elongated).
- Poro-colpate. Combination of alternating colpi and pores around the equator.

According to their position on the pollen, apertures can be (cf. Fig. 3.10):

- Zonoaperture. Located at the equator.

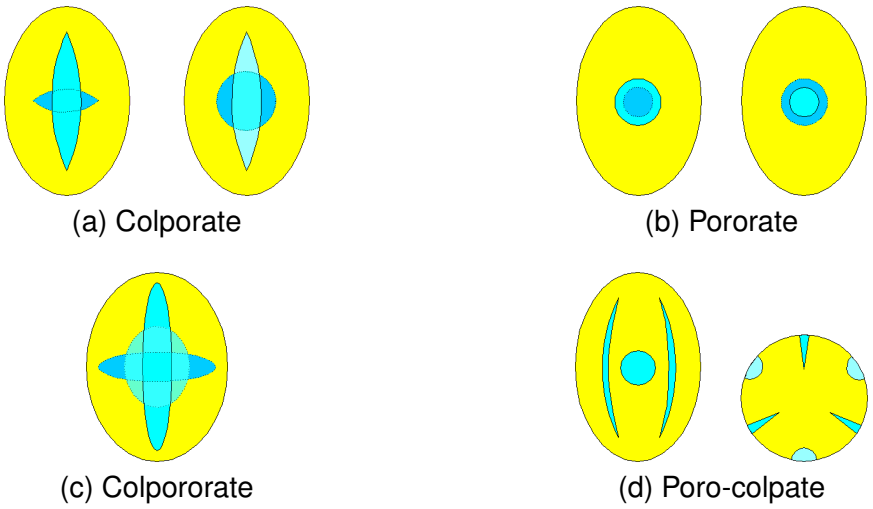


Figure 3.9: Combinations of apertures types. Images from [Punt et al., 2007].

- Anaaperture. Located at the distal face.
- Cataaperture. Located at the proximal face.

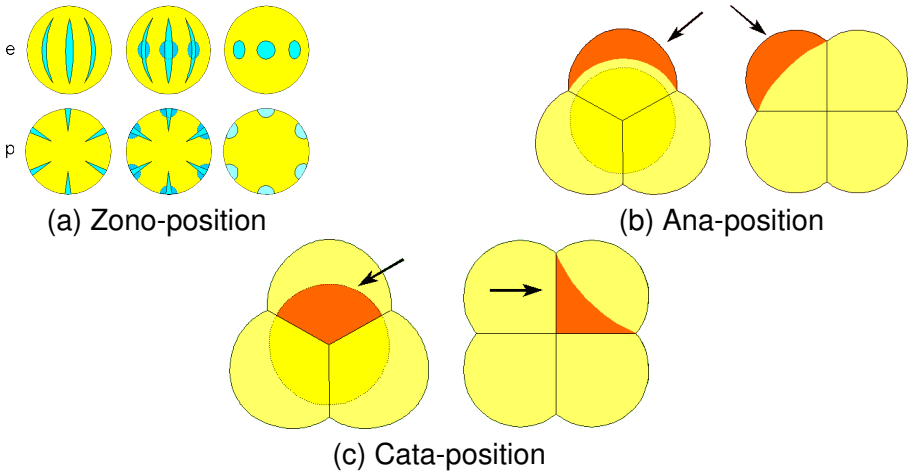


Figure 3.10: Positions where apertures can be located. Images from [Punt et al., 2007].

Additional structures that modify the morphology of the apertures are (cf. Fig. 3.11):

- Annulus. Surrounding border of a pore with different thick or shape to the rest of the exine.
- Arcus. Thickening of the sexine connecting apertures in a sweeping curve from.
- Oncus. Lens-shaped structure beneath the aperture.

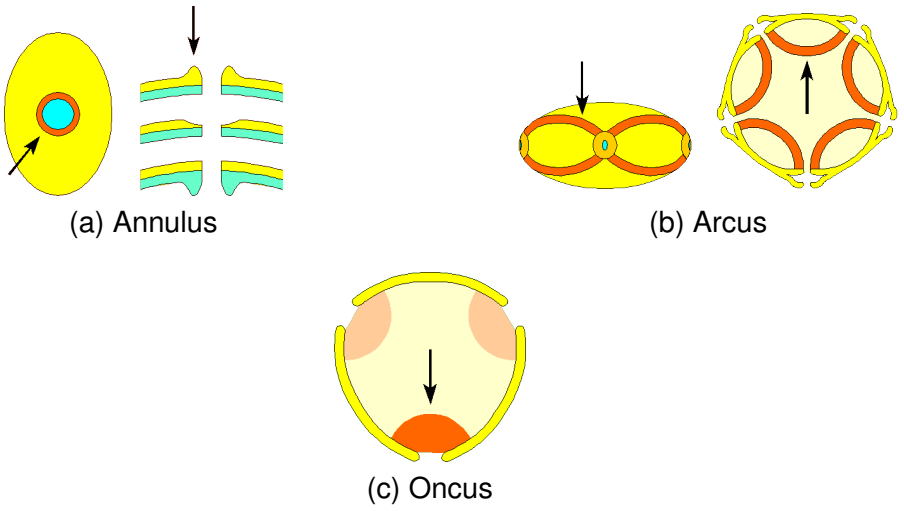


Figure 3.11: Type of modifying structures of apertures. Images from [Punt et al., 2007].

3.3.2/ PHENOLOGICAL INFORMATION

The presence of pollen in the environment obeys the flowering season of each specific plant. These cycles are studied by the phenology. Historical pollen count records allow the determination of a calendar of the expected concentration of pollen along the year. Recognition strategies have employed differences of concentration among the taxa to reduce the range of possible taxa in a determined period. This is especially important for similar taxa with different flowering season. Estimations of the concentration calendar vary depending on the year and the region. A recent concentration calendar for Germany is depicted in Fig. 3.12.

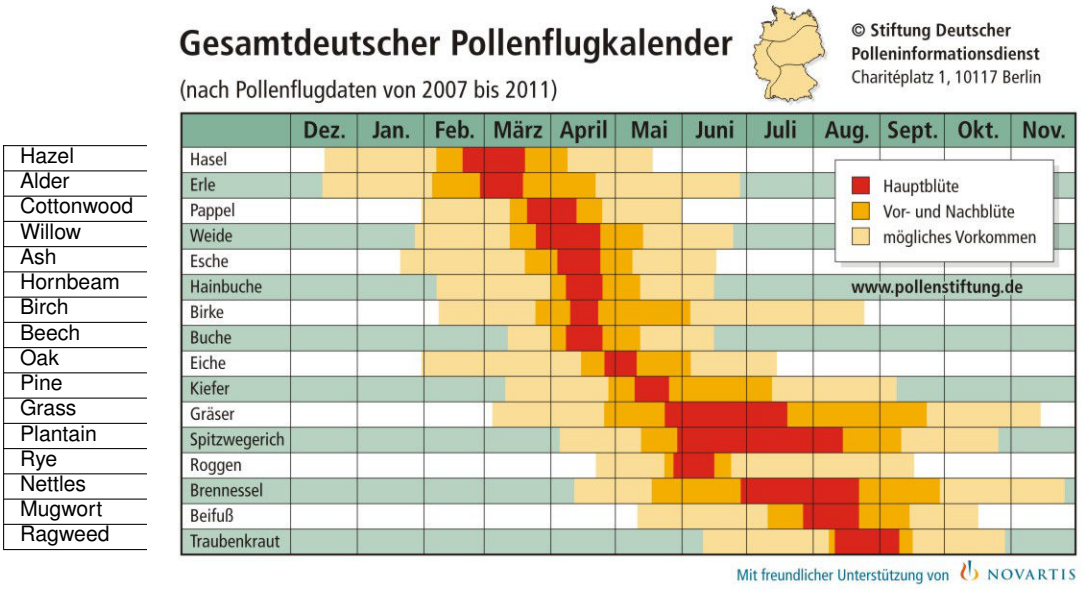


Figure 3.12: Pollen concentration calendar estimated by the PID from 2007 to 2011 for the entire region of Germany. Red periods indicate high concentration. Amber periods indicate moderate concentration. Light yellow periods indicate low concentration. Green periods indicate no concentration [3].

3.3.3/ STUDIED POLLEN

In Table 3.3, palynological information of the five studied pollen is gathered. For the case of statistical measures, such as the case of size and concentration periods, global estimations for Germany are considered. When the proposed recognition method is applied to a particular environment, the measures should be adjusted to the specific time period and geographic region for a more precise analysis.

[3] Due to copyright, modification of the original image is not allowed. Pertinent translations are provided. Image from www.pollenstifung.de.

Table 3.3: Palynological characteristics of the studied taxa taken from [Buchner and Weber, 2014], except size from the Austrian Pollen Information System[4], and the concentration calendar from Fig. 3.12.

Pollen taxon	Alder	Birch	Hazel	Mugwort	Grass
Family	Betulaceae	Betulaceae	Betulaceae	Asteraceae	Poaceae / Gramineae
Genus	Alnus	Betula	Corylus	Artemisia	Many
Shape	spheroidal, oblate	spheroidal	spheroidal	spheroidal, prolate	spheroidal, prolate
Polar view outline	polygonal	circular, irregular	circular, triangular	circular, lobate	circular irregular
Size [μm]	20 - 25	21 - 25	25 - 31	20 - 24	34 - 36
Ornamentation	scabrate	psilate	psilate	echinate	psilate
Aperture number	4 - 5	3	3	3	1
Aperture type	colpate	porate	porate	colporate	ulcerate
Aperture peculiarity	arcus	annulus, oncus, operculum,	annulus, oncus, operculum,	sunken when dry	annulus, operculum
Concentration calendar	January to March	March to May	January to March	June to September	April to October

[4] www.pollenwarndienst.at

4

Feature Extraction for Pollen Representation

The analysis of the palynological characteristics taxa from Table 3.3 reveals very similar characteristics among taxa, for example inside the family *Betulaceae* where differences are very subtle. This inter-class similarity is difficult to overcome even for expert palynologists, leaving the class decision sometimes only to probabilities based on the concentration calendar. Furthermore, the variability of appearances among individuals of the same taxon (intra-class) is a difficulty always associated to biological entities (*cf.* Fig. 3.3).

The present study proposes the description of pollen based on features extracted from the image of the particle in an image-based pattern recognition framework. This process involves the preparation of images, segmentation, feature extraction, feature selection and classification. The required set of features must be able to overcome the aforementioned inter-class similarity and the intra-class variability.

In the first part of this chapter, an overview of the classification system is given, followed by the description of the methods used for localization and segmentation. The second part explores the different characteristic groups by providing detailed description of the analyzed features. In particular, the analysis of apertures is left to an additional chapter because they require a special treatment and for which a novel method is developed. Finally, a brief comment regarding to the use of color features is provided.

The deep analysis of the characteristic groups is conducted in Ch. 6, in order to determine the most relevant. The selected subsets will be used in a global classification scheme, whose results are given in the same chapter.

4.1/ Overview of the Classification System

After the preparation of slides with collected pollen and their digitization, the first steps consist of the localization of particles on the slide, and their validation as potential pollen. Images are pre-processed in order to normalize and enhance visual characteristics. A segmentation algorithm is applied in order to detect the surrounding contour of the particle, and thus, to detect the shape of the particle and demarcate the particle and the background. In these steps, debris is eliminated as well as pollen with odd characteristics.

The single-taxon dataset is employed to extract individual 2D snippets of pollen particles

for each pollen type. Due to the random spread of the particles, their orientation and their position along the depth of the medium are not fixed. The result is a representative sampling of the variability of orientations and focal planes at which the particle can be found.

The feature set employed for the analysis of pollen is divided in groups according to their computational base and to the descriptive knowledge used by palynologists for the characterization of shape, ornamentation, and apertures. These characteristics groups are General Shape Features (GSF), Elliptic Fourier Descriptors (EFD), Texture, and Apertures. Benefiting from advances on object recognition techniques, statistical features that enable the description of the pollen characteristics are applied to each particle. However, features do not represent exactly the same measures that palynologist do. Instead they provide useful information in form of numeric data that usually are disregarded by humans eyes due to their difficulty of being observed and computed.

The pattern recognition process consists of two phases. First, for the training phase, the whole set of proposed features are extracted from the contour of the pollen particles. A feature selection method is applied to each characteristic group in order to determine the most discriminative subset according to its performance in a classification task. Then, a classifier is trained to recognize the taxa based on the selected subset of features. The scheme of the system in this phase is represented in Fig. 4.1.

For the testing phase, the relevant features are extracted from a subset of the pollen particle and are input to the learnt classifier model in order to evaluate the performance of the proposed process. In this manner, the robustness of the method to unseen particles is assessed.

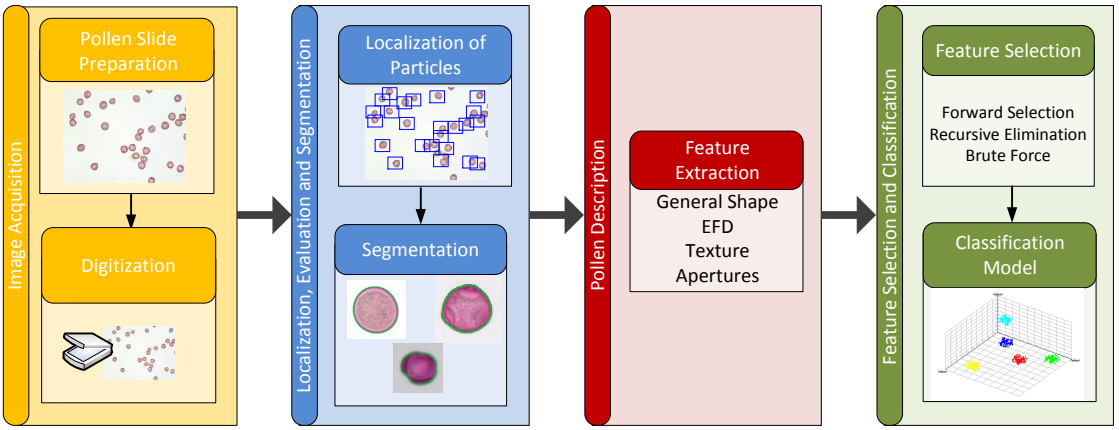


Figure 4.1: Overview of the proposed statistical system in the training phase where the feature selection and construction of the classification model take place.

4.2/ POLLEN EXTRACTION

In a real world application, airborne pollen is cluttered with debris; particles that are mostly different enough to allow a simple discrimination (*cf.* Sec. 3.1). However, due to their high incidence in the airborne samples (debris/pollen ratio has been estimated to be about

0.90), it is important to reject them before the classification process in order to alleviate the processing load of subsequent processes and to reduce variability on the data.

First, a lower resolution image of the slide is employed to reduce the size of the processed area. At this step, detailed information is not necessary. The original image is subsampled by a ratio of 0.25 and converted to gray levels.

A noise reduction is carried out by applying a median filter on the image. This operation eliminates speckle debris and noise without the risk of eliminating pollen particles. Then, the whole image is inverted, which leaves the particles on bright levels while the background on dark ones. This makes the image easy to interpret and enables the subsequent methods easy to apply.

With the purpose of increasing the separability between the particles and the background, the contrast is enhanced. In this step, the borders of the particle are better bounded but also the appearance of the particle gray-level pattern is modified, although at this point it is not important and changes will be reverted later. The method employed for enhancing the contrast is the linear stretching. Figure 4.2a shows the histogram of gray values of the inverted image. Most of the pixels belong to the background, depicted as black and dark values. The contrast of a specific range in the histogram is stretched by applying a linear transformation. The range of interest (where the particles lie) is chosen using the mean of the histogram (peak of the background pixels) at the lower side and the maximum existing gray level on the upper side. The resulting histogram is shown in Fig. 4.2b.

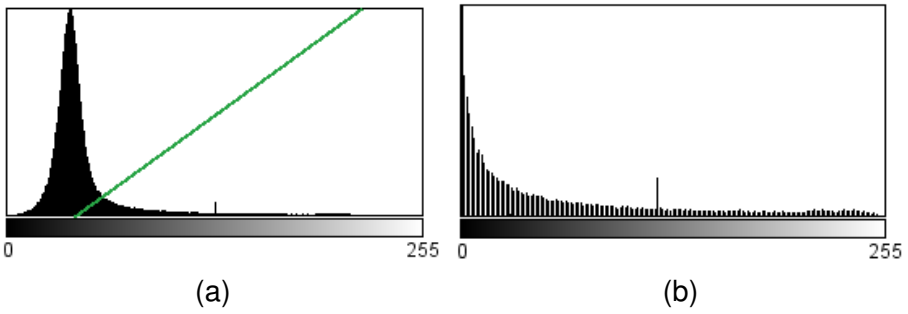


Figure 4.2: Contrast enhancement of the slide: (a) histogram of the slide showing high frequencies for background pixels. Linear stretching of the range of interest is shown as a green line and (b) Histogram of the slide after the application of the contrast enhancement.

It is possible then to identify clearly the particles on the processed slides. The next step is to apply a thresholding in order to binarize the image. Due to the clear borders of the particles, the threshold value is not critical and it is chosen empirically. The binarization shows a very close representation of the particle shapes, even though some internal patterns of the particle remain in black level as holes. With this operation, the background also becomes black.

Opening morphological operation is applied to the image with the purpose of eliminate both, thin particles and image artifacts due to fuzzy borders or debris stuck to the pollen. Opening consists of applying two basic morphological operations consecutively, erosion and dilation. The kernel size of three pixels was chosen as good trade-off between eliminating undesired objects and avoiding major modifications to the pollen shape.

Finally, image snippets are taken for each of the individual particles found on the pre-processed image. This is performed by considering each white blob in the image as an individual particle. The blob detection is done by estimating the blobs contour using the Suzuki's method [Suzuki and Abe, 1985]. The method employs a border-following strategy on the binary image to detect the blobs. Then, it classifies them into outer or hole border types. Finally, it builds the topological structure of the borders using a surroundedness relation search. For the case of pollen slides, only outermost borders, corresponding to individual particles, are of interest. For each particle, a bounding box is fitted to the blob and its position and size are kept. This information is used for the extraction of snippets from the original-scale slide.

The localization method performs excellent most of the time, and only fails when pollen particles are not well dyed and lack of contrast. An example of the application of the method is depicted in Fig. 4.3.

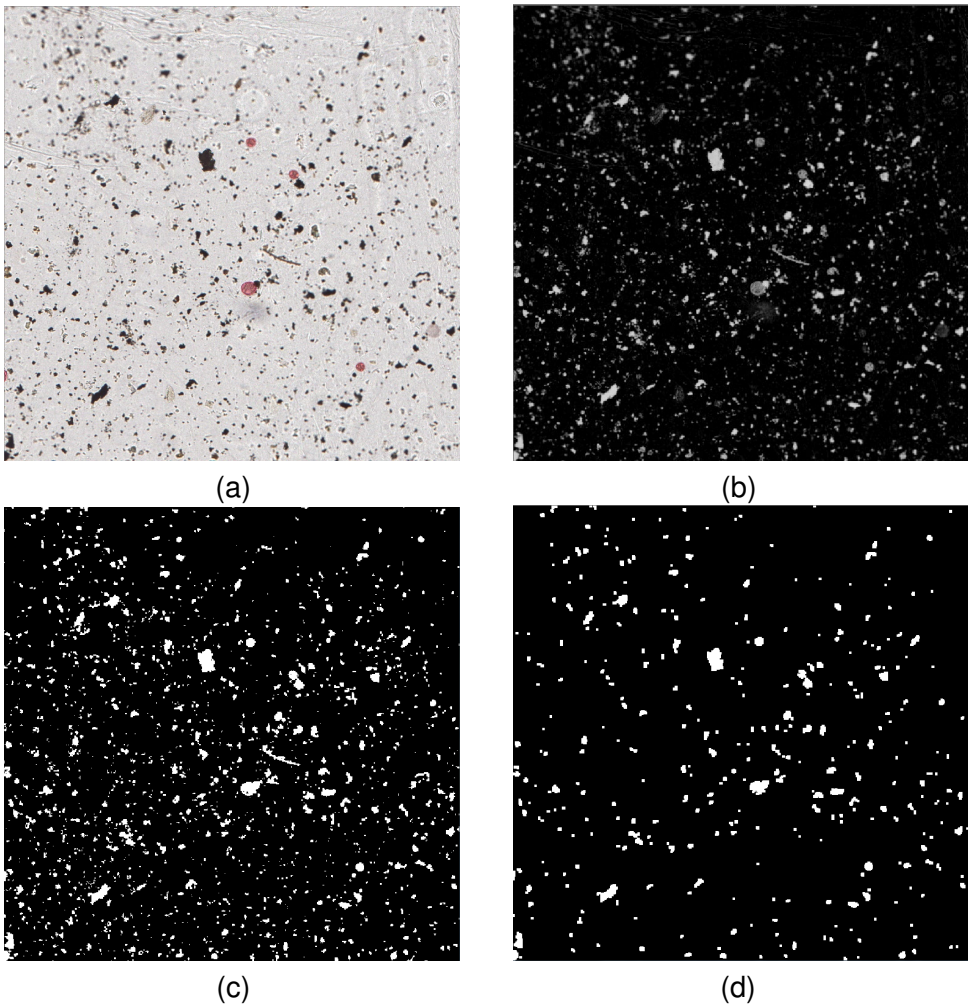


Figure 4.3: Example of localization of particles on a slide: (a) original slide section, (b) inverted slide after downscale, noise reduction and contrast enhancement, (c) slide after application of threshold, and (d) particle blobs after opening morphological operation.

Before extracting the particles snippets, an initial pre-computation is needed. Due to possible variations on illumination during the scanning process among different acqui-

sitions and on the transparency of slide medium, the background of the slide needs to be compensated. The overall mean RGB value of the slide image can be regarded as the background level since the proportion of area containing particles compared to the free area is very low. Therefore, the color of the particles does not affect considerably the mean. This value is subtracted only until the snippets are extracted once they are confirmed to be potential pollen in order to avoid the processing on rejected particles.

The information from the localization step (from the downscaled image) is converted to get the position of the snippet at the original scale. Then, individual snippets are extracted from the original color slide image.

Evaluation of potential pollen snippets is performed in order to keep in the database only those that are highly likely to be pollen particles. The localization process employs for this task the most discriminant characteristics and yet the simplest to detect. The near-circular shape, similar pink color and bounded size are the foundation of the processing.

First, the size of the snippet is evaluated. Typical diameter of the studied pollen is in the range of 10 to 50 μm. Translated to pixels, the range is from 40 to 200 for a resolution of 0.25 μm/pixel and from 29 to 143 for a resolution of 0.35 μm/pixel. Therefore, snippets with size out of this range in either width or height are rejected. This first test is fast and simple to perform.

The second test is the pink color of the pollen after the dyeing. Since different pollen types respond different to the dye and acquire different color tonalities, the reference color should be adequate for all the pollen types and dyeing condition.

The color of the particles is extracted from a centered internal box of one fifth the size of the snippet in order to ensuring that the sample covers the particle. Hue and Saturation values of pixels are considered as the color representation. Value component from the HSV space is ignored in order to reduce luminance variations. A particular Hue and Saturation range of the datasets is estimated by analyzing pixel-wisely the color of the pollen particle.

Results showed that 90% of the pollen pixels present values of Hue in the range of 324-360° and of Saturation in the range of 0-50%. Pixel outliers, which are mainly due to light diffractions or defective dyeing, are dropped for the computation of these ranges. The outliers are easily identified because their Hue-Saturation values are very different from the typical pink color.

The evaluation of unknown particles considers the rate of pixels having the typical pollen color. The color of every pixel is evaluated whether it belongs to the pollen Hue and Saturation range and the proportion of positives is computed with respect to the total of sampled pixels. Particles having a ratio below the threshold of 0.25 are rejected. Although this threshold seems low, debris has very low or null pink content. On the other hand, the low threshold allows accounting for the variation of pink tonality among different pollen taxa and dyeing methods.

The third filter is the circular shape detection of the pollen. Modelling pollen as circular objects helps to discriminate pollen from irregular shapes. Hough circle detection, which is derived originally from the Hough line detection, is the technique employed for this purpose [Duda and Hart, 1972]. The principle of the Hough circle detection is the supposition that every pixel in the image can be part of a locus of circles parametrized by

$$(x - x_0)^2 + (y - y_0)^2 = r^2, \tag{4.1}$$

where x_0, y_0 and r are the center and radius of the circle. Then, based on the particular image shape patterns, the belief of actual circles on the image is strengthened by a voting process.

A 3D array of parameters (x_0, y_0, r) is employed to accumulate the vote of every pixel for all possible parametrized circles that fit. Finally, local maxima of the 3D array are interpreted as the most probable circles are regarded as detected circles.

It is evident that the computation of a 3D array for all possible circles for all the pixels in the image is very impracticable. The Hough gradient is the method applied for the computational simplification of the Hough circle detector [Kimme et al., 1975]. First, the number of evaluated pixels is reduced by considering the gradient image. After detecting edges the Canny filter, first-order Sobel gradients are computed on both directions, rows and columns. A second simplification is the consideration for the array of parameters only those circles whose center is pointed by the gradient direction of the evaluated pixel. Furthermore, only potential centers are kept in a simplified 2D array of parameters (x_0, y_0) regardless the radius. The employment of a 2D array reduced considerably the computation load of the estimation of local maxima. The estimated centers are kept if enough pixels support it with the same radius length and it is apart enough from other centers. Finally, the process is also constrained to search on a specific radius range according to the expected pollen size, which further reduces the number of detected circles.

The application of the process yields to a list of potential circular shapes inside of the particle snippet. If at least one circle of diameter greater than 15 μm is detected, the snippet is finally kept in the database of pollen, otherwise it is rejected. Figure 4.4 shows examples of the successful evaluation of circular shapes. After this point, differences on the illumination of the background are compensated on the remaining dataset by adjusting the illumination of the snippets so that the background is always at the same white level.

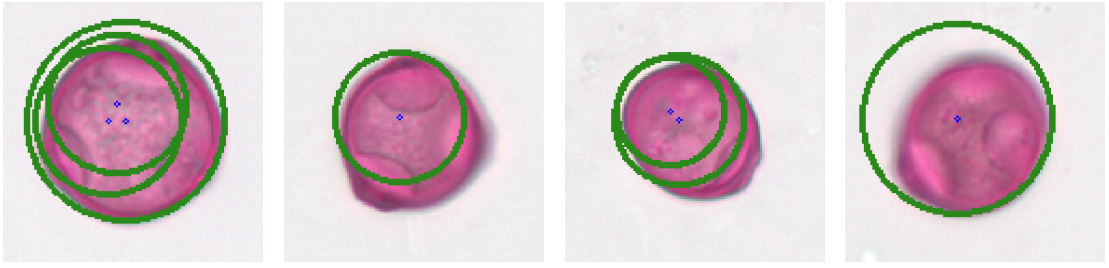


Figure 4.4: Detection of circular shapes on pollen samples using Hough circle detection.

For single-taxon datasets, the empirical rate of potential pollen is about 74% of the total of 535 extracted particles. The rest 26% of the particles are rejected mainly due to the out-of-focus snippet, bigger size than allowed, irregular shape, debris, and stuck particles. In particular, lack of focus prevents a correct estimation of the particle contour and of the features (if the particle were not rejected), and therefore, it is advantageous to reject blurred images in this step. In this strategy, a strict rejection threshold is preferred in order to have a clean dataset.

In the case of airborne datasets, the rejection rate reaches 98% (489 particles) of the 500 extracted objects, mainly due to an important content of debris, with an estimation

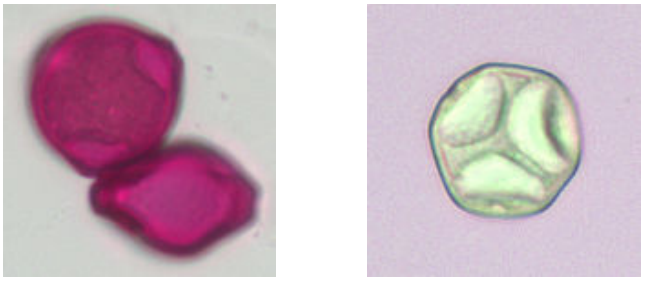
of about 92% (461 particles). The rest of the rejections are due to out-of-focus snippets, irregular shape, and stuck particles. Only 6% (28 particles) were rejected because they were pollen with one of the aforementioned sub-optimal conditions.

Because pollen is not perfectly round, some flexibility must be allowed. This flexibility could hardly cause debris to be accepted since most are not circular shaped. Additionally, the color and size filters reduce this possibility. However, when two particles are stuck, the circle detector could either missed them or detect them as a single bigger particle. In such case, the size filter would fail too.

This effect can be avoided on laboratory tests with the adequate slide preparation. Stuck particles were found rarely in the single-taxon datasets, reaching only 1.0% of the rejected particles. For the airborne datasets, the effect depends only on the concentration of particles and debris in the air. Therefore, stuck particles cannot be controlled in this kind of samples. Nevertheless, it was found empirically that stuck particles represent about 3.5% of the rejected particles of the airborne dataset.

Another case of failure of the pollen evaluation is when pollen has deficient dyeing and it does not acquire the typical pink color. This problem is avoidable during the slide preparation and can be reduced completely with the correct application of a dyeing protocol.

Examples of both cases are depicted in Fig. 4.5.



(a) (b)

Figure 4.5: Examples of not detected pollen: (a) Two stuck particles seem to be a bigger particle, and (b) deficient dyeing.

4.3/ SEGMENTATION

Once individual snippets of the potential pollen have been extracted, segmentation of the detailed contour of the particles is performed. This step is very important for the subsequent extraction of shape-based features because many of them are based on this contour. For the rest of features, a good estimation is enough for their correct computation.

An efficient method for the segmentation of this kind of particles is based on their characteristic pink color. As in the localization step, the color of pollen can be characterized by a range in HSV color space. Due to the required precision, the range of hue and saturation should be chosen more precisely according to the expected color of the current slides. The original size of the image is employed. Different ranges are obtained if the conditions of the image acquisition or if the slide dyeing are changed.

By assigning to each pixel of the snippet a value of 1 if they belong to the pink color range and 0 otherwise, a binary mask is obtained which defines the particle shape. It is convenient to apply a median filter in order to smooth the mask to eliminate single speckle pixels around the contour that do not contribute to the pollen shape although they are in the pink range. These pixels could be result of light diffractions around the particle. A small kernel of five pixels is enough.

Fuzzy regions on the border between the particle and the background caused by lack of focus can also cause irregularities on the contour when in reality the border is even. In addition to the smoothing at the previous step, closing operation, which consists in dilation and erosion of the particle mask, is applied to reduce this effect. However, this operation can also destroy actual irregular shaped regions on the contour, as in the case of some mugwort views (*cf.* Fig. 4.6d).

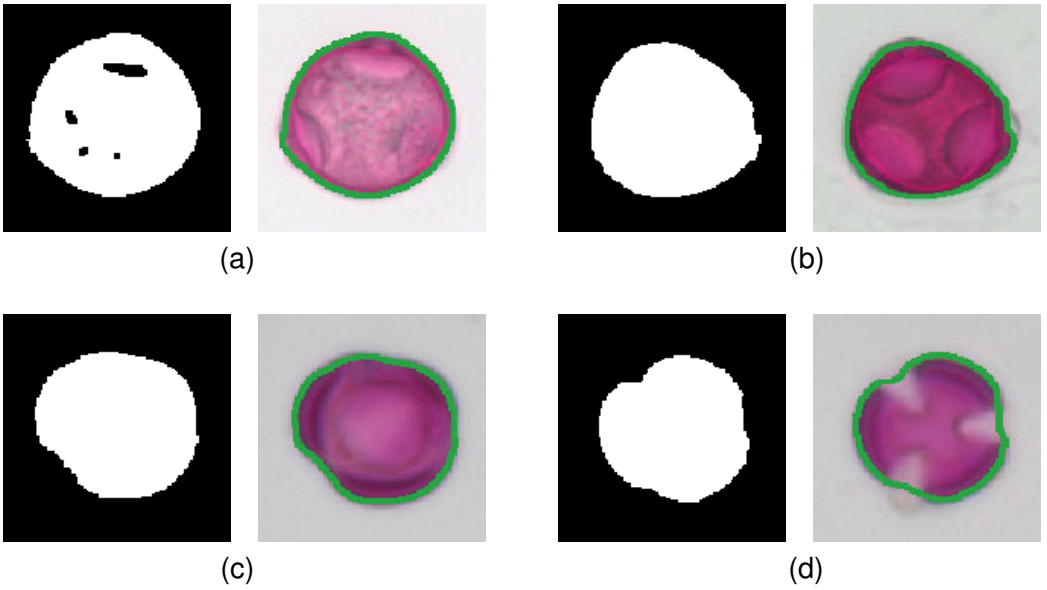


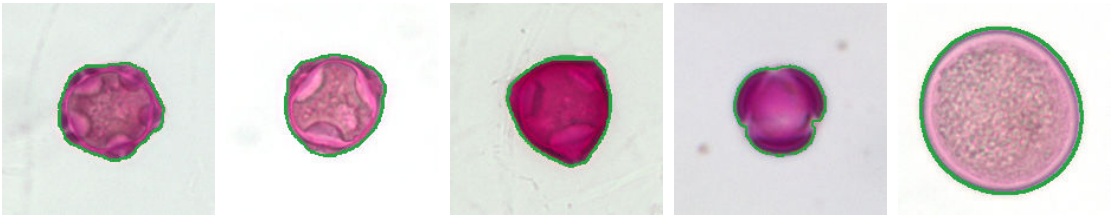
Figure 4.6: Examples of correct segmented pollen using color-based approach are shown in (a)-(c). On the left side of each example, the final binary mask, from which the contour is extracted, is shown. On the right size, the contour is depicted in green. Imprecise segmentation in (d) is due to a fuzzy and irregular border.

The color-based approach is very useful and precise when a standard method of acquisition and dyeing can be established. Examples can be seen in Fig. 4.6. However, because of the different possible ranges and deficient dyed slides, the method requires parameter adjustment for every new analyzed slide. This adjustment is not practical for the variability of the current dataset. In such cases, an automatic method was employed similar to Chen *et al.* [Chen et al., 2006]. The segmentation method employs Otsu's thresholding because it is nonparametric, unsupervised and fast [Otsu, 1975].

The method consists of searching the threshold value that splits a gray level image into foreground (particle pixels) and background such that the intra-class variance is minimized. Therefore, a different threshold is estimated for each snippet. It was found experimentally that the method can determine the correct threshold even under variation of light intensity due to the scanning process, as long as a certain intensity difference between the background and the particle exists (*cf.* Fig. 4.7). Furthermore, it is not necessary that

the particle is well pink dyed.

After the binarization, it is easy to trace the contour using the Suzuki's method as explained in Sec. 4.2. Performance of this segmentation method is excellent provided that the source image exhibits a good contrast and there is no debris stuck to the pollen.



Alder Birch Hazel Mugwort Grass

Figure 4.7: Example of the five pollen taxa of interest and their segmentation using Otsu's automatic method. The detected contour in green is enhanced for visualization.

4.4/ Feature Extraction

4.4.1/ General Shape Features

Recalling section 3.3.1, information about the morphology of the pollen has been much resorted for the recognition of taxa. In this framework, previous image-based approaches have proven that certain shape features are suitable to characterize pollen for classification purposes [Chen et al., 2006; Rodríguez-Damián et al., 2003]. The type of General Shape Features (GSF) employed until now, are statistics that measure a property of the shape at a time. For example, recalling Table 3.2, the P/E ratio can be considered as a GSF. Equivalent GSFs need to be employed if pollen images are considered as 2D slices of a volume, from a specific focal plane. Therefore, volumetric measures such as P/E ratio are not possible in this kind of data. The shape considered for the computation of GSFs is based on the contour shape extracted by segmentation.

Most of the classical GSF are based on well-known primary measures *perimeter* and *area*. Because they are dependent on the shape size, the resulting classical GSF are combined in such manner to make them size invariant, and to measure solely a single characteristic of the shape. The proposed classical GSF are summarized in Table 4.1 and described in detail below [Da Fontoura Da Costa and Cesar Jr, 2000].

Table 4.1: List of proposed classical general shape features.

Classical shape features

Feature	Short name	Features	Short name
Perimeter	P	Radius dispersion	rdis
Area	A	Ratio ζmax/ζmin	ratio1
Roundness	R	Ratio ζmax/ζmean	ratio2
O'Higgins Undulation	U	Ratio ζmin/ζmean	ratio3
Complexity f	cxf	2n Euclidean norm	2eN

***Perimeter* (P)**

It is computed as the length of the surrounding contour.

***Area* (A)**

It is the estimated surface bounded by the surrounding contour.

With focus on describing the complexity of the shape the following GSF are considered.

***Roundness* (R)**

Also known as *thinness ratio*, it is computed as:

$$R = 4\pi \frac{A}{P^2} \tag{4.2}$$

This feature indirectly measures the elongation and irregularities on the contour (undulation). In such irregular shapes, the perimeter will increase respect to the area reducing the ratio towards 0. When the shape is a perfect circle, $R = 1$.

***O'Higgins Undulation* (U)**

A similar undulation ratio related to the *roundness* and also dependent on the *area* and the *perimeter* is defined as [O'Higgins, 1997]:

$$U = \frac{P - \sqrt{P^2 - 4\pi A}}{P + \sqrt{P^2 - 4\pi A}} \tag{4.3}$$

***Complexity f* (cxf)**

This is an alternative measure of the complexity of a shape. Considering the minimum distance between a point r inside the shape to the shape contour c, the mean distance (β) of all the inner points can be computed as:

$$\beta = \frac{1}{N} \sum d(r, c), \tag{4.4}$$

where $d(a, b)$ is the minimum distance function between a point a and a set of points b. Then, the *complexity measure f* is defined as:

$$cxf = \frac{A}{\beta^2} \tag{4.5}$$

Due to the computational expense of estimating β for all the points, the distance transform can be employed. The distance transform is defined as the assignation to each external point P of a shape S the value corresponding to the minimum Euclidean distance between them. The Chamfer distance algorithm proposed by Borgefors was employed [Borgefors, 1986]. This algorithm is widely used because of its efficient computation and yet reasonably accuracy. Although modern algorithms, for example the one in Ref. [Saito and Toriwaki, 1994], focus on improving efficiency and accuracy, the performance of the Borgefors'algorithm is reasonable sufficient considering that pollen shape contours are little abrupt, and that values are eventually averaged.

A transformation grid (mask) of 3x3 pixels was applied to compute the distance, and the distance measured as the cost ($= 1$) of horizontal, vertical or diagonal shifts. Therefore, β

can be computed by admitting only the distance transform value of the inner points.

***Radius dispersion* (rdis)**

The computation of this feature is associated to the centroid of the shape. The geometric centroid M is computed by the average of all the points N that form the contour $c(n)$ as follows:

$$M = \frac{\sum_{n=0}^{N-1} c(n)}{N}. \tag{4.6}$$

The *radius dispersion* is defined as the standard deviation of the distances between M and all the points in $c(n)$. Let $\zeta(n)$ be the Euclidean distance between $c(n)$ and M:

$$\zeta(n) = \sqrt{(c(n) - M)^2}. \tag{4.7}$$

Then, the *radius dispersion*, which measures the variability of the shape with respect to the centroid, is computed as given below. The case of perfect circle yields to $rdis$ equals to 0.

$$rdis = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(\zeta(n) - \overline{\zeta})^2}. \tag{4.8}$$

Similarly, the following three features employ also the distance $\zeta(n)$ to estimate some ratios.

***Ratio* ζ*max*/ζ*min* (ratio1)**

Following the measures with respect to the centroid, three additional ratios can be estimated, with the additional advantage of lacking of dependence on the shape size. $Ratio1$ is defined as:

$$ratio1 = \frac{max(\zeta(n))}{min(\zeta(n))} \tag{4.9}$$

This ratio compares the extreme points with respect to the centroid. For a perfect circle, $ratio1$ is equal to 1.

***Ratio* ζ*max*/ζ*mean* (ratio2)**

Similarly, $ratio2$ is defined as:

$$ratio2 = \frac{max(\zeta(n))}{mean(\zeta(n))} \tag{4.10}$$

This ratio indicates how far is the maximum extreme point with respect to the rest of the shape, and it is sensitive to lobes. For a perfect circle, $ratio2$ is equal to 1.

***Ratio* ζ*min*/ζ*mean* (ratio3)**

Finally, $ratio3$ is defined as:

$$ratio3 = \frac{min(\zeta(n))}{mean(\zeta(n))} \tag{4.11}$$

Similarly to ratio2, this ratio indicates how far is the minimum extreme point with respect to the rest of the shape, and it is sensitive to cavities. For a perfect circle, $ratio3$ is equal to 1.

2n Euclidean norm **(2eN)**

Consider the 2n-vector representation of a shape contour c by concatenating the x, y values for each point $\vec{c} = [(x_1, y_1); (x_2, y_2); ...; (x_n, y_n)]$. Vector norms can be applied to shapes in order to extract a measure of its size. Applying the 2n Euclidean norm to $\vec{c}$, normalizing to translation with respect to the centroid coordinates m_1, m_2, and normalizing to the number of points N, the norm is computed as:

$$2eN = \|\vec{c}\| = \sqrt{\frac{\sum_{i=1}^{N}(x_i - m_1)^2 + \sum_{i=1}^{N}(y_i - m_2)^2}{2N}} \tag{4.12}$$

Together with the *area* and the *perimeter*, the *2n Euclidean norm* are features that depend on the actual size of the particle, which in practice means that they depend on the resolution of the input image. Although this condition could seem rather restrictive, actually they are employed as an indirect measure of the pollen size. For this reason, they are dropped out from the analysis of shape representation and considered as size-related features.

Axes Ratio **(EF_ratio)**

The elongation of the shape is also important to describe the pollen outline. An additional method to describe the shape is the fitting of an ellipse to the shape. The proposed features summarized in Table. 4.2 reflect directly the elongation of the shape and the deviation of the shape from a smooth outline (irregularities) in an independent manner. The fitting algorithm employs the optimization proposed in [Ahn et al., 2001]. The fitting method is based on finding the coordinate description of a point on the ellipse that corresponds to the shape point, given the orthogonal distance as mean to find the shortest path. The corresponding ellipse points are given by the ellipse parameters and the shape. Then, the Jacobian matrix on the ellipse point and a least-square iterative fitting is performed to minimize the square sum of the error (geometric fitting). From the resulting ellipse, the following features are computed:

Direct measure of the elongation of the shape is given by the ratio between axes: $EF\_ratio = a/b$, where a and b are the long and short axes of the fitted ellipse respectively.

Given the fitting error as the distance between c points and the fitted ellipse, statistics related to irregularities are computed as the ***mean of the error*** **(EF_mean),** ***standard deviation of the error*** **(EF_std)** and the ***Root Mean Square (RMS) error*** **(EF_rms)**. EF_rms is associated to EF_mean and EF_std, and the information could be redundant. However, the RMS is included in the evaluation because it could be a simple form of representing the fitting error instead of both the mean and the standard deviation.

Table 4.2: List of proposed ellipse fitting general shape features.

Ellipse fitting features			
Feature	Short name	Features	Short name
Axes Ratio	EF_ratio	Mean error	EF_mean
RMS error	EF_rms	Std Dev of the error	EF_std

4.4.2/ Elliptic Fourier Descriptor

It is not possible to represent all the variability of the pollen shape only with single-value GSF due to its high complexity. The focal plane can cut the particle in many different positions with result of different particle outlines for the same taxon. For this reason, the GSF are complemented with a powerful and detailed description of the shape outline.

Elliptic Fourier Analysis (EFA) is a technique based on Fourier decomposition of the shape outline treated as the combination of a pair of signals, which in the discrete form are represented by the shape coordinates. The shape description can be split into components ranging from rough to fine. These components can be studied later (for example, by applying PCA) to determine which contain more morphological information about the shape, and in which occur most of the variation.

Due to the flexibility of selecting the number of harmonics of the expansion, the quality of the representation can be adjusted to the desired level. EFA has been already applied to the study of shape patterns with biological intra-class variation: EFA has enabled the understanding of the major shape variation and in some cases the classification of plant leaves [Neto et al., 2006], petals [Yoshioka et al., 2004], fruits [Menesatti et al., 2008; Goto et al., 2005; Costa et al., 2009], cereal grains [Mebatsion et al., 2012; Williams et al., 2013], fossils [Crampton, 1995], roots [Iwata et al., 1998], and mosquito wings [Rohlf and Archie, 1984].

The computation of the descriptor based on EFA is as follows: The continuous curve $c(t)$ defining the contour of a shape can be expressed as the complex combination of the curves representing the behavior in direction of the x and y axes as:

$$c(t) = x(t) + jy(t) \tag{4.13}$$

where t is the displacement along the contour, $c(t)$ has a period $T = 2\pi$ (length enough to trace all the points of the contour) and a fundamental frequency $\omega = 1$. The coefficients of the Fourier expansion are given by

$$c_k = c_{kx} + jc_{ky} \tag{4.14}$$

where k represents each of the harmonics. This can be expressed also as:

$$c_k = A_k - jB_k \quad \text{and} \quad c_{-k} = A_k + jB_k \tag{4.15}$$

where:

$$A_k = \frac{a_{xk} + ja_{yk}}{2} \quad \text{and} \quad B_k = \frac{b_{xk} + jb_{yk}}{2} \tag{4.16}$$

The terms in eq. 4.16 can be approximated by the following discrete expressions because the contour $c(t)$ consists of N discrete points.

$$a_{xk} = \frac{2}{N}\sum_{i=1}^{N} x_i cos(k\omega i\tau) \quad \text{and} \quad b_{xk} = \frac{2}{N}\sum_{i=1}^{N} x_i sin(k\omega i\tau) \tag{4.17}$$

$$a_{yk} = \frac{2}{N}\sum_{i=1}^{N} y_i cos(k\omega i\tau) \quad \text{and} \quad b_{yk} = \frac{2}{N}\sum_{i=1}^{N} y_i sin(k\omega i\tau) \tag{4.18}$$

Therefore, the complex exponential representation of the curve $c(t)$ based on the coefficients summation is given by

$$c(t) = c_0 + \sum_{k=1}^{\infty}(A_k - jB_k)e^{jk\omega t} + \sum_{k=-\infty}^{-1}(A_k + jB_k)e^{jk\omega t} \tag{4.19}$$

The four fundamental terms in eq. 4.17 and 4.18 allow the representation of the shape up to a level of accuracy given the number of considered harmonics. This representation can be understood as an elliptic phasor that points to the current shape point. For each harmonic k, the fundamental terms represent a different ellipse with changing locus. At a determined instant t, the locus of each ellipse is defined where the preceding phasor points at. The entire location of the shape is determined by a_{x0} and a_{y0}.

Nixon and Aguado proposed a shape descriptor by combining the terms in eq. 4.16 as defined in eq. 4.20 [Nixon, 2008]. The resulting Elliptic Fourier Descriptor (EFD) is invariant to translation, rotation and scale, and it does not involve negative frequencies.

$$EFD_k = \frac{|A_k|}{|A_1|} + \frac{|B_k|}{|B_1|} \tag{4.20}$$

For the implementation of the EFDs on pollen particles, contours have been scaled so that they all have the same length. This adjustment allows all shapes to have the same number of points, which is equivalent to equalize the sampling rate among the shapes. This operation does not affect the representation of the particle since both features are scale invariant.

Aforementioned analyses on different biological individuals have considered from seven to 50 harmonics for the representation of the contour. This representation achieves reasonable quality considering that, in most of the cases, the individuals belong to similar taxonomic classification and the observations are fixed to a single viewing angle. In the present study, the first 150 harmonics are computed from the particle contours after leaving out the first two harmonics, which are uninformative due to the normalization. This difference in the number of harmonics is intended for a more complex classification task involving more taxa with variable viewing angle. Moreover, after a relevance analysis presented in Sec. 6.2.2, some harmonics could be discarded.

4.4.3/ TEXTURE FEATURES

Exact description of the physical texture patterns formed in the pollen wall is too complex to be defined by a fixed model because random biological variation is present on the distribution of the constituting elements. The texture sensed by the camera and reflected on the digital image is result of the distribution of pixel values in a neighborhood. Those values are product of the interaction of the pollen ornamentation, the pollen cytoplasm, the intensity of the light, the image resolution, the lens magnification, the focal plane, and the interference from out-of-focus planes.

All these sources of variation prevent the pollen texture from being described just by templates or geometric models. Therefore, statistical methods take into account this variation and enable the computation of measures that are suitable to represent different texture classes. Examples of texture pattern of pollen are shown in Fig. 4.8.

(a) (b) (c) (d) (e)

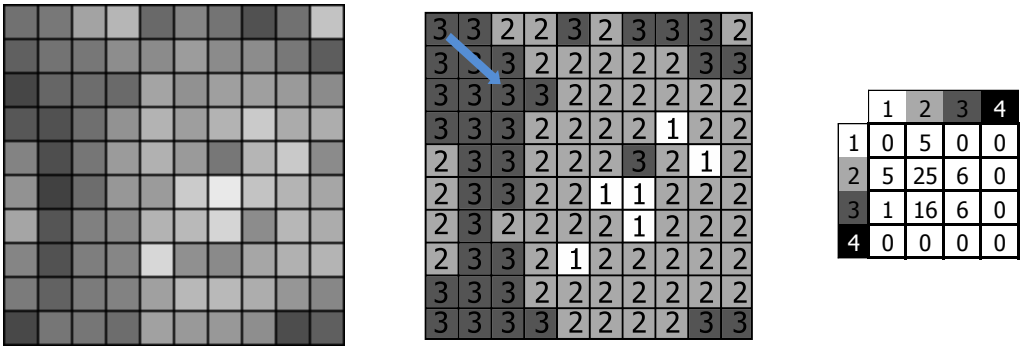
Figure 4.8: Example of texture patterns of (a) alder, (b) birch, (c) hazel, (d) mugwort, and (e) grass. Images are taken from the central part of the particle, shown in gray levels and with contrast enhanced in order to improve visualization and to remark differences among types.

The Gray Level Co-occurrence Matrix (GLCM) is a well-known statistical method to interpret the texture patterns on gray-level images. The GLCM computation is based on the spatial relation of pairs of pixels. Haralick suggested some statistics from the GLCM as texture descriptors [Haralick et al., 1973]. The Haralick's features have been widely employed in tasks of texture-based classification, for example, of tumors [Zulpe and Pawar, 2012], aerial landscapes [Mhangara and Odindi, 2013], vegetation [Chabrier et al., 2012], and wood [Tou et al., 2009].

The GLCM is computed as the second-order histogram of the frequency of the combination of gray values from a pair of pixels, given a separation offset. The offset can be expressed by a horizontal and a vertical offset in pixels, which can be converted to a diagonal distance and an angle. To each gray value corresponds a column and a row in the GLCM.

In order to reduce the number of possible combinations, and thus the size of the GLCM, the gray-level image is quantized to N_g gray levels (GLCM becomes N_gxN_g). This reduction not only reduces the computational expense, but also, generalizes the texture patterns to avoid specialization. This means that after quantization, similar pattern will have very similar GLCM (and texture features), enabling a correct classification. The computation of the GLCM is depicted on Fig. 4.9.

The GLCM is sensitive to monotonic gray-level changes because it employs the absolute pixel values for quantization. An important step to solve it is the pre-processing of the

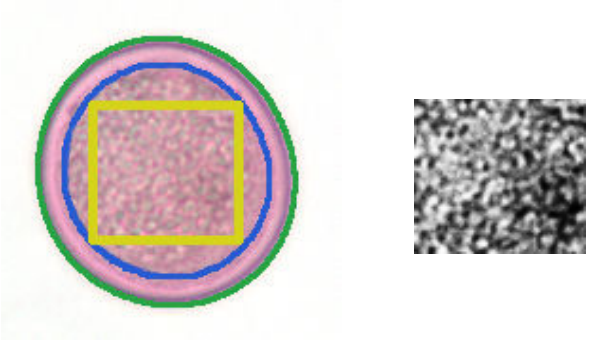


(a) (b) (c)

Figure 4.9: The original gray-level image (a) is quantized to four gray levels resulting in the image (b). The offset indicated by the blue arrow in (b) (in this case [2,2]) is employed for comparing the pairs of pixels. The second order histogram of frequencies for the given offset is shown in (c). Because there are not black pixels in the image, the column and row of the matrix for level 4 are blank.

image against illumination variation, which is achieved by equalizing the image histogram. The goal of this technique is to stretch the histogram in order to improve the contrast by redistributing the intensities. It is achieved by a transformation of the frequencies of the histogram such that its cumulative distributive function becomes linear in the whole range of gray values.

Representative texture patterns are formed on the inner part of the particle of the pollen image. For this reason, the image is sub-sampled by an image patch that fits a shrunk version of the contour. In this manner, the border of the particle and the background are not examined. This process is depicted in Fig. 4.10. To avoid differences on the frequencies of the GLCM due the variation of the patch size, the GLCM is normalized to 1 to obtain relative frequencies.



(a) (b)

Figure 4.10: (a) Original contour (green) is shrunk by erosion of the original shape. A sampling box (yellow) is fitted to the modified contour (blue). Texture sample (b) after equalization of the histogram.

Given the normalized GLCM, $p(i, j)$, auxiliary definitions are expressed as:

$$p_x(j) = \sum_i p(i, j), \tag{4.21}$$

$$p_y(i) = \sum_j p(i, j), \tag{4.22}$$

$$p_{x-y}(k) = \sum_{\substack{i=1 \\ |i-j|=k}}^{N_g} \sum_{j=1}^{N_g} p(i, j), \tag{4.23}$$

$$p_{x+y}(k) = \sum_{\substack{i=1 \\ i+j=k}}^{N_g} \sum_{j=1}^{N_g} p(i, j), \tag{4.24}$$

with $\mu_x, \mu_y, \mu, \sigma_x, \sigma_y$ and σ being the means and standard deviations of p_x,p_y and p. Eleven Haralick's features are computed as follows:

Angular second moment:

$$f_1 = \sum_i \sum_j p(i, j)^2 \tag{4.25}$$

Contrast:

$$f_2 = \sum_{k=0}^{N_g-1} k^2 p_{x-y}(k) \tag{4.26}$$

Correlation:

$$f_3 = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{4.27}$$

Sum of squares (variance):

$$f_4 = \sum_i \sum_j (i - \mu)^2 p(i, j) \tag{4.28}$$

Inverse difference moment:

$$f_5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j) \tag{4.29}$$

Sum average:

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i) \tag{4.30}$$

Sum variance:

$$f_7 = \sum_{i=2}^{2N_g} (i - f_8)^2 p_{x+y}(i) \tag{4.31}$$

Sum entropy:

$$f_8 = -\sum_{i=2}^{2N_g} p_{x+y}(i) log \left\{ p_{x+y}(i) \right\} \tag{4.32}$$

Entropy:

$$f_9 = -\sum_{i} \sum_{j} p(i,j) log \left\{ p(i,j) \right\} \tag{4.33}$$

Difference variance:

$$f_{10} = \text{variance of } p_{x-y} \tag{4.34}$$

Difference entropy

$$f_{11} = -\sum_{i=0}^{N_g - 1} p_{x-y}(i) log \left\{ p_{x-y}(i) \right\} \tag{4.35}$$

For the analysis of pollen, texture descriptors are computed considering 25 different pixel offsets ranging from one to seven pixels in distance, equivalent to the range from 0.5 to 2.9 μm. Angle offsets are -90°, -45°, 0°, 45°, and 90°. An illustration of the offsets around the reference pixel is shown in Fig. 4.11 The rest of the orientations that complete the circumference around the reference pixel, is already contained on the information of the opposite angle (for example, information at 180° is contained at 0°). The quantization level N_g is four.

The combination of 25 offsets with eleven Haralick's measures results in 275 texture features. They are used as representative of the pollen texture in the analysis stage.

4.4.4/ Color Features

The study of color as discriminant characteristic for the taxa has been limited. The subjective perception of color by the palynologist during the pollen count would yield to inaccurate results without the help of the correct technology. In addition, dyeing alters the natural color of the particle, being dependent on the substance, the amount, and the employed technique.

Image-based approaches have employed little color information for classification and its suitability is not convincing yet (*cf.* Sec. 2.2.5). However, alternative studies of the spectrum of the pollen suggest subtle difference among taxa (*cf.* Sec. 2.2.9). Further color and spectrum analysis requires strict control of the luminosity in the acquisition of the image and the application of a standardized dyeing method.

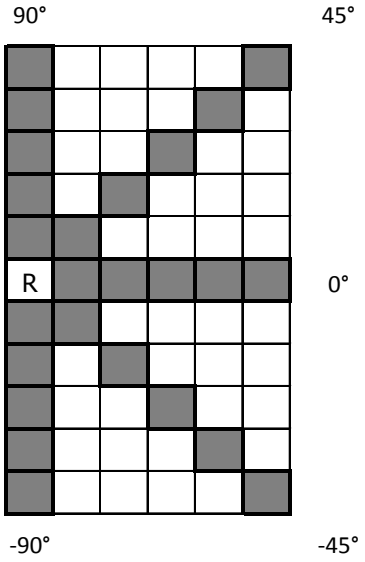


Figure 4.11: Topology of the employed offsets with respect to the reference pixel R. Pixels used in the computation of the Haralick's features are shown in gray.

In the current study, the color of the pollen is employed exclusively for helping the localization and segmentation of the particles, although no color features or characterization is applied, mainly because the dyeing on the pollen is subject to variation.

4.4.5/ Application of the Characteristic Groups to Pollen Representation

The features presented in this chapter have the goal of describing different pollen characteristics. Pollen external morphology is represented by two groups of features according to the level of detail. First, a set of features describes the complexity of the pollen contour in a very general manner, for example, by measures that express the irregulaity or the roundness of the particle. Some of the these features are related to the size of the particle.

The second set of features is based on a more sophisticated representation of the shape of the pollen contour, which considered detailed changes. This representation employs EFD harmonics, which account for frequency contributions.

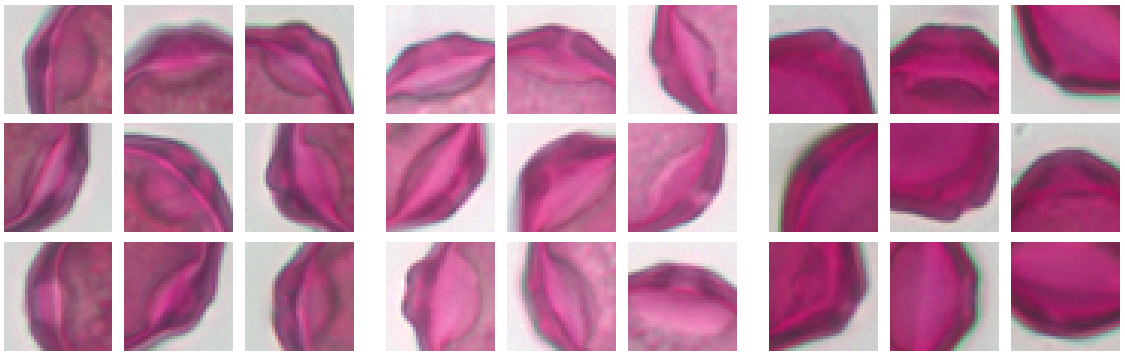
The ornamentation of the pollen is represented by a set of Haralick's measures on the texture patterns. In order to account for their complexity, patterns are described in different combinations of orientation and distance offsets.

Due to the complexity of the morphology of the pollen particles, it is not expected that the characteristic groups are able of a great taxon classification when applied individually because they are focused on just a single characteristic. However, the combination of these groups enables the synergic contribution of their strengths, and at the same time, mutually counteract their weaknesses. Therefore, it seems reasonable to expect that the classification result of the joint application of the characteristic groups to be importantly better than individual schemes. The complete analysis of the features groups is conducted in Ch. 6.

5

Aperture Detector

Determination of the type and number of apertures is important for the recognition of pollen taxa (*cf.* Sec. 3.3.1 and Tab. 3.3). This task represents an important challenge since genera of the same family are often similar, and likewise, the apertures. Intra-genus variation is due to the different morphological appearances of the apertures of the same taxon. Fig. 5.1 exemplifies the diversity of appearances of apertures from the three genus of the family *Betulaceae*, on which this investigation is focused: alder, birch, and hazel. The case of apertures of mugwort is not considered in this approach because they are in unfavorable position for observation. The size of the apertures of grass is very small, which makes difficult a robust representation using the proposed method. Moreover, their small size makes them difficult to show up, preventing an adequate dataset size. Therefore, apertures of grass are also not considered for the detector.



(a) alder (b) birch (c) hazel

Figure 5.1: Examples of the diversity of appearances of apertures of (a) alder, (b) birch, and (c) hazel pollen. Note that rotation is also present in the dataset.

Due to this complexity, the need of a flexible method arises in order to overcome the variability due to changes of the viewing angle. In this chapter, an approach is proposed for the detection, localization and counting of apertures of multiple pollen types. The method has the advantage of eliminating the need of designing new algorithms for each type. Therefore, the growing of the detector with more apertures is straight forward.

Overview

The proposed method based the description of the apertures on their content of primitive images. The collection of these primitive images builds up a visual codebook following a similar procedure to the Bag-of-Words (BOW) approach. In this method, primitive images

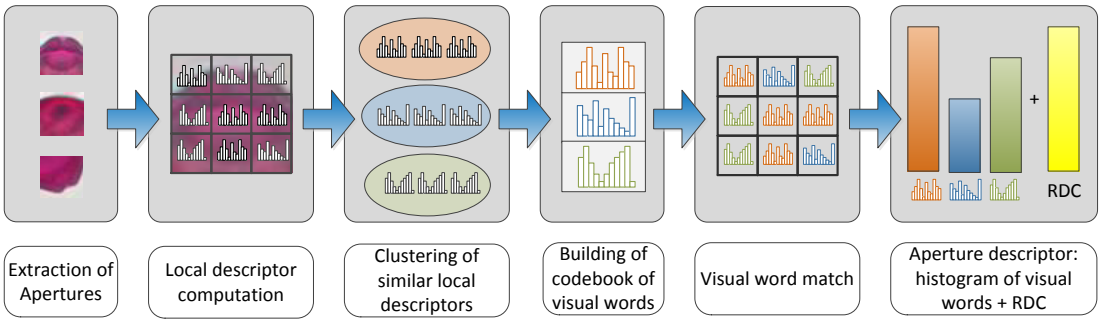


Figure 5.2: Computation of the aperture descriptor: Local descriptors are computed on a dense grid of aperture regions and clustered to create the codebook of visual words, which are depicted by a different color. Finally, apertures are described by the histogram of visual words combined with the relative distance to the centroid (ρ) as spatial information.

are represented by a local descriptor and clustered according to their similarity. Therefore, any region of the image can be described in terms of a unique combination of visual words from the codebook. Moreover, spatial information of each word is added to the representation. The overall process is represented in Fig. 5.2. Based on this representation, a classification model is created to discriminate aperture regions from the rest of the particle. This methodology is applied to different pollen taxa to create individual classifiers oriented to the variety of apertures.

For the detection of apertures on unseen particles, multiple regions inside the particle are classified as positive or negative using the proposed model. A confidence map, which indicates the likelihood of the presence of apertures, is created from the pixel-wise averaged votes of the regions. The position and number of the apertures are computed by means of a local-maxima search on the map. Finally, the definitive detected aperture list is the aggregation of results of the individual classification for all the aperture types. This evaluation process is shown in Fig. 5.3.

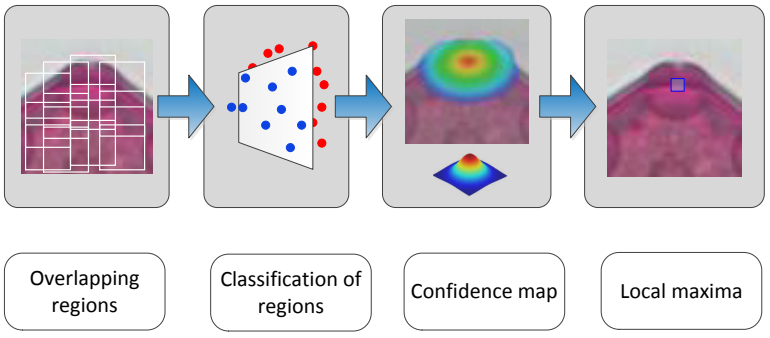


Figure 5.3: Detection scheme of an aperture on an unseen particle: Multiple overlapping regions are evaluated and classified to create the confidence map from which the aperture position is estimated, marked as a blue square.

The rest of this chapter describes in detail the proposed aperture detector. First, an

introduction to the BOW method is presented, followed by the description of each of the BOW steps. Then, an extensive evaluation is conducted. Finally, conclusions of the overall method and evaluation are drawn.

5.1/ Application of the Bag-of-Words Approach

The foundation of the aperture recognition strategy is the description of visual patterns present in the images in terms of primitive subimages. This idea originated from the Bag-of-Words (BOW) strategy employed for text categorization, in which sentences are characterized by a set of words, from a vocabulary or codebook. Then, the frequency of occurrence of the words in the text allows their classification [Joachims, 1998]. The approach has been adapted to visual object recognition tasks, where the words are replaced by image patches, taking the name of bag of keypoints or bag of features. This adaptation is not straight forward since the definition of the visual words from image patches and the construction of the codebook can be defined in different ways. Choices have to be made for the descriptor of the patches and for the similarity measure when building the codebook. Some successful BOW strategies are described in Ref. [Csurka et al., 2004; Lazebnik et al., 2006], and improvements on these choices have been also proposed in Ref. [López-Sastre et al., 2011; Wang et al., 2013]. However, to our knowledge, the proposed method is the first application of BOW to microscopic objects.

The essential procedure of the BOW method employed in our approach is, in essence, the same proposed by Csurka *et al.* [Csurka et al., 2004], consisting of the four following steps:

- Sampling and description of the image patches.
- Creation of the codebook and matching the visual words to the patches.
- Representation of the image by the visual words occurrence.
- Classification of the objects employing their feature vector representation.

The following sections detail these steps and their adaptation to the aperture detector.

5.2/ Sampling Image Patches

In the first step, image patches of the aperture are represented by a local descriptor. Typically, this sampling process in BOW is performed with the help of the detection of salient keypoints. With this technique, a reduced number of patches with relevant information are sampled. An alternative is to increase the number of sampled patches by using a regular sampling with a grid, which increases the number of carried information from the object. Although this technique conveys an intense computation, experiments have proven that the increment of information yields to better classification results [Nowak et al., 2006]. For the aperture detector, the latter strategy is chosen. Due to the reduced size of the aperture regions, less than 40 pixels, even a dense sampling that covers the whole region is computationally affordable.

A square grid of square patches is extracted exactly over the aperture region as depicted in the example of Fig. 5.4. In this manner, most of the patches are informative. The size of the sampled region is result of two choices: the number of patches by side N, and the size p of each patch, in pixels by side. The resulting region has $N \times p$ pixels by side. From now on, the region configuration will be denoted by these values as region N-p.

Figure 5.4: Example of a grid 4x4 employed for the computation of local descriptors. The patch size is a square of 8 pixels.

5.3/ Description of the Image Patches

The following step is the representation of visual patterns from the patches. Here, a local descriptor is employed to extract the visual information. Additionally, spatial information of the patch is taken into account. It is important to consider the ability of the descriptor to represent shape or patterns. Comprehensive surveys of local descriptors can be found in Ref. [Mikolajczyk and Schmid, 2005; Goyal and Walia, 2014; Nanni et al., 2012]. In the current study, the Local Binary Patterns (LBP) operator is employed as the local descriptor due to its efficiency on the representation of shapes and patterns. More detail about this descriptor is given in the next section.

Since the pollen is scanned at even resolution and magnification, it is not mandatory to count with complete scale invariance. Moreover, the aperture size is regular for the same pollen taxon and differences are too small. Rotation robustness is important because apertures can appear at any orientation on the pollen image. This property is achieved by employing samples in multiple orientations and by the rotation invariance of LBP.

5.3.1/ Local Binary Patterns

In this section, the theoretical background of LBP is provided. Considerations for the extraction of visual information from the image patches are also indicated. The LBP descriptor is a simple operator that efficiently encapsulates the spatial pixel arrangement in the vicinity of a central pixel [Ojala et al., 1996]. LBP has been widely employed with success for the recognition of faces, task that encounters problems like object deformations and light intensity changes [HuoRong et al., 2013; Huang et al., 2011]. Application in the medical/biological domain seems suitable, based on the description of micro-patterns, for example in Ref. [Lladó et al., 2009; Unay and Ekin, 2008]. It is particular suitable when the characteristic feature appears at small scales like in the case of the aperture patches. Due to their success, numerous variants have appeared with the goal of improving their representation power [Nanni et al., 2012].

Haar-like features have been the straightforward alternative employed with similar purposes, for example, face recognition [Rodriguez, 2006]. This feature computes neighborhood influence based on five masks derived from Haar wavelets. The advantage of LBP over Haar-like features is the invariance to gray-level variations. Furthermore, less LBP descriptors are needed to achieve comparable discrimination as evaluated in Ref. [Rodriguez, 2006].

The grey level of a reference pixel located in (*x,y*) is compared one-to-one to the levels of neighbor locations. The basic case is the comparison of the central pixel with its 8 immediate neighbors. The result of the comparison is marked with a value of 0 if the reference pixel has a greater gray level than the corresponding neighbor, and 1 otherwise. Then, a binary code is constructed by counterclockwise concatenating the individual comparison results as bits.

The relative comparison strategy, in contrast to absolute pixel levels, makes the operator invariant to monotonic gray-level changes on the image. More complex approaches extend the compared neighborhood to a radius r from the central pixel in order to look for bigger pixels patterns. They also allow for the limitation of the number p of surrounding pixels, in order to discard very similar comparisons and to keep a compact representation of the final code. With these two variable parameters, a LBP code is denoted as $LBP_{p,r}$. A graphical representation of the computation of a LBP code is shown in Fig. 5.5.

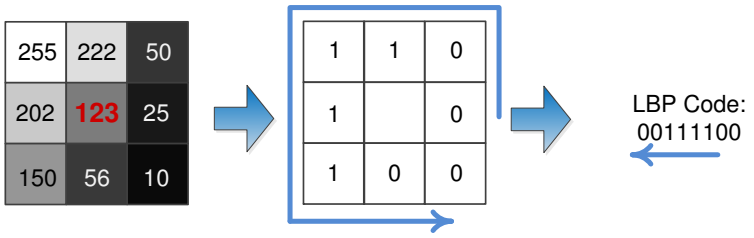


Figure 5.5: Example of computation of the LBP code from a single pixel. (a) Values of the gray levels of the eight neighbors are shown overlapped to each pixel location. The value of the reference pixel is highlighted in red at the center. (b) Individual binary results of the comparison to the reference value. (c) Final concatenated LBP code.

When the area is rotated, the position of the compared pixels in the neighborhood is shifted, altering the order in which the comparison bits are concatenated. This yields to different codes for similar pixel arrangements. To solve this inconvenience, Ojala *et al.* identified 36 local binary patterns in a eight circular neighborhood, which are independent to pixel circular shifts [Ojala et al., 2000]. They assigned a single code for each pattern. Beside the reduction of the possible LBP codes, this technique provides additional rotation invariance.

In a further simplification, Ojala *et al.* defined a set of uniform patterns with individual LBP codes in which uniformity is defined by the maximum number of transitions among bits (0/1 change) on the LBP code. The rest of non-uniform patterns were grouped together in a single code. Figure 5.6 shows the rotation invariant uniform patterns with maximum two transitions, $r = 1$ and $p = 8$ denoted as $LBP^{riu2}_{8,1}$. LBP are reduced to ten possible codes (nine codes for the patterns 00000000, 00000001, 00000011, 00000111, 00001111, 00011111, 00111111, 01111111, 11111111 and an additional code for the

non-uniform patterns). Ojala *et al.* proved that the rotation-invariant uniform patterns of an image are still effective since they represent on average 90% of the textural information [Ojala et al., 2000].

The $LBP^{riu2}_{8,1}$ operator is selected for the aperture detector due to its simplicity and compact representation in addition to its tested effectiveness. The 8-pixel neighborhood is suitable to test different patch sizes. The selection of the uniform patterns and the computation of the LBP depend strongly on scale of the image. Since the image resolution is fixed at the image acquisition stage, the scale invariance of the LBP becomes of little relevance and a fixed neighborhood can be chosen.

The LBP descriptor of each of the sampled patches is obtained by the application of the LBP operator on all the pixels of the patch, except for the one-pixel border. Then, the histogram of the resulting LBP codes is computed and normalized yielding to the LBP descriptor vector.

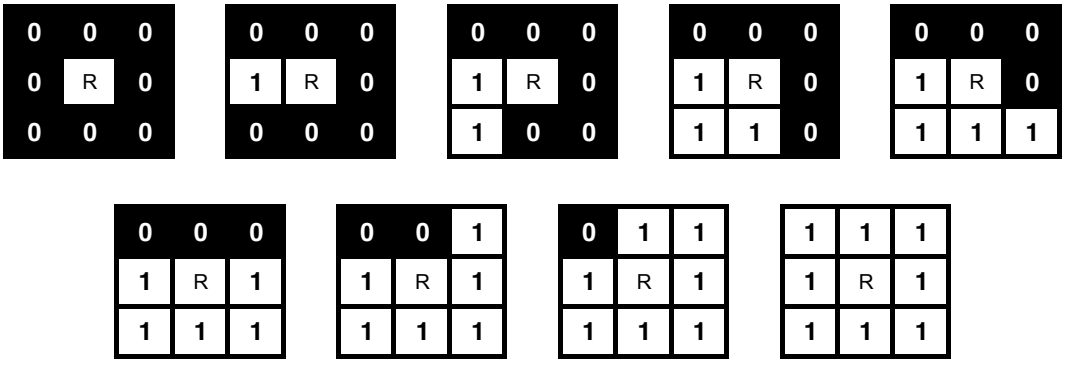


Figure 5.6: The nine uniform pixel patterns proposed by Ojala *et al.* [Ojala et al., 2000], corresponding to $LBP^{riu2}_{8,1}$. Black and white neighbor pixels correspond to bit values 0 and 1 after LBP comparison to the reference pixel R.

5.3.2/ Spatial Information of the Patch

An additional improvement to the LBP descriptor is the incorporation of spatial information of the patch. Experiments showed that adding spatial information can improve the BOW performance [Zhang and Mayo, 2010; Cao et al., 2010]. In the proposed method, a numerical value, which indicates the position in the grid at which each patch is extracted, is appended at the end of its LBP descriptor vector. Considering that grid cells are evenly spaced and sized, a fractional value ranging from 0 to 1 is assigned to each, with a step equal to *1/number of cells*. The assignation of the spatial code to a region 4x4 is illustrated in Fig. 5.7

The extended LBP descriptor vector is used in the next section to create the codebook of visual words. In a similar manner, this local-feature vector is employed to match a testing patch to the already created visual words.

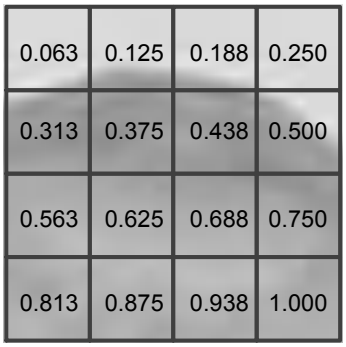


Figure 5.7: Example of the spatial codes assigned to each position of a grid 4x4.

5.4/ CREATION OF THE CODEBOOK OF VISUAL WORDS

The codebook of visual words is the collection of basic visual patterns, or primitives, that constitute the aperture appearance. This section explains how this codebook is built from the LBP description vector of the image patches. In the general case, a regular visual pattern can be simply represented by the same local-feature vector because the pattern remains unchanged and has always a constant vector. However, in the case of apertures, where visual patterns present a natural variability, local-feature vectors are similar but not exactly the same. Hence, different local-feature vectors are innumerable and are not suitable individually for representation in the codebook.

The BOW solution is to group those similar local-feature vectors into a single visual word, which is indeed, representative of all the associated patterns. Usually BOW employs k-means for this purpose, a square-error partitioning technique for clustering and quantizing local-feature vectors. Starting from k clusters centers in the feature space, an iterative process assigns each vector to the closest clusters while minimizing the within-cluster square error. Then, it updates the centers of the clusters with the new vector members.

The method does not indicate how to select the number k of clusters, although there are additional techniques that search a good estimation [Pelleg et al., 2000; Tibshirani et al., 2001]. However, the flexibility of choosing k is convenient for the aperture detector because enables the selection of the codebook size that best recognizes apertures, according to the classification performance. Experiments to choose the best k are conducted during the experimental evaluation.

The procedure to create the codebook is to collect all the local-feature vectors of the patches of the aperture from the training dataset as input of the clustering method and to create k automatic clusters using the Euclidean distance as the similarity measure. Therefore, the centers of the cluster are the representation of the visual words in the codebook. For evaluation of an unknown patch, its local-feature vector is matched to the closest visual word. Consequently, all the patches of an aperture region are assigned to a known visual word.

5.5/ Representation of the Aperture

In the third step of the BOW procedure, the representation of the aperture is built from the content of visual words of the created codebook. First, the patches of the aperture image are matched to the closest visual word according to the k-mean cluster model. Then, the histogram of the occurrence of the visual words for each aperture is computed. The number of bins of the histogram is equal to the size of the codebook k. Finally, the histogram is normalized to show relative frequencies. The advantage of this representation is that visually similar apertures will have a similar histogram distribution.

Specific aperture appearances are commonly in regular positions on the particle, with respect to the centroid, for example on the border. This information helps to reject morphologically similar regions on the particle that cannot be considered apertures because of their atypical position. In order to estimate the position of the regions and yet keeping comparability among taxa of different size, a relative measure is employed. The relative distance (ρ) between the center of the region and the centroid of the particle is estimated. In this manner, the region situated at the centroid receives a value of 0, and a region at the outermost point of the particle receives a value of 1. The centroid of the particle is taken from the shape features described in Sec. 4.3.1. Figure 5.8 shows an example of ρ for different positions.

The definitive aperture descriptor vector is built by concatenating the vector of histogram frequencies and the distance ρ. With this unique descriptor, the classification of regions is possible in order to discriminate the aperture regions from the rest. The classification of regions supports the determination of an aperture in a particular area of the particle as described in the following section.

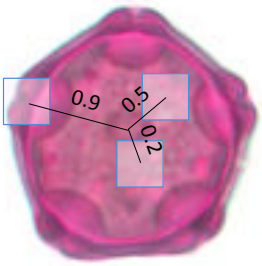


Figure 5.8: Example of some regions at different positions inside the particle and its distance ρ to the centroid.

5.6/ Individual Classifiers

In this section, the strategy of classification of regions is presented, explaining the importance of splitting the task into individual classifiers. One of the important shortcomings of the approaches on aperture detection described in Ch. 2, which is not solved until now, is the flexibility to integrate multiple aperture types and appearances into a single method. This characteristic makes possible to incorporate eventually new appearances to the detector in a simple training process without altering already the learned appear-

ance models and employing the same description method. For this reason, the aperture detector is split into individual classifiers, each dedicated to one appearance. At the end, the result of all classifiers is combined for a final conclusion of the detected apertures.

An individual classifier is trained independently from the others, meaning that internal parameters can be adjusted optimally for each case. When a new appearance is needed to be added to the detector, an new individual classifier is built and integrated to the final phase of the detector.

The final detected apertures are the sum of all detected apertures individually, and therefore, the detection problem presents the optimal substructure property. In consequence, the global optimal performance can be obtained by the individual optimal performance of the classifiers. In this regard, an individual classifier is for each pollen taxon whose apertures are required to be recognized. Hence, single classification tasks cope with the intra-class variation.

Due to the possibility of variation on the morphology of the aperture even of the same taxon, the selected classification strategy has to be robust enough to cope with this complexity. Support Vector Machine (SVM) method has been chosen due to its proven efficacy in binary classification problems in image processing [Byun and Lee, 2002]. SVM strategy is based on mapping data into higher dimensions where linear separation is possible.

Using the training dataset, a binary SVM model is created for each individual classifier in the fourth step of the BOW. Positive examples correspond to regions that contain an aperture. On the contrary, negative examples correspond to regions that do not contain an aperture.

The Radial Basis Function (RBF) is employed as the kernel function, enabling the non-linear transformation of the original feature space into a higher space. This kernel uses the parameter γ for bounding its range. Additionally, the SVM employs the parameter C as a weight of penalization of the misclassifications during the process of finding the optimal separating hyperplane. The parameter values are chosen in a grid search in the experimental part.

Cross-validation technique is employed for reducing the overfitting effect and evaluating the generalization capacity of a statistical model. This assessment technique splits the dataset into x subsets of the same size. Then, the statistical model, in this case the classifier, is trained with all the subsets except one, which is left for testing and computing the performance. Repeating the process while alternating the testing subset, the definitive performance is given by the average of all the iterations.

In order to enable the classifier to better discriminate negative examples, examples of regions from the whole database, are considered for training. Regions that do not contain apertures are much varied and numerous than those showing an aperture since, most of the pollen area is not an aperture. This yields to an unbalanced dataset. This problem is addressed by random sub-sampling the negative examples in order to reduce their number and yet keeping the representation of the variability of the negative class.

For training, positive and negative regions are directly extracted from the pollen image because their position within the image is known. With this precise information, training regions are accurately centered on the aperture. However, for testing, the unknown orientation and position of the aperture make difficult to sample box-shaped regions that exactly fit the aperture. Therefore, it is important to incorporate additional regions that are

not exactly centered on the aperture to the training set, so that partially occluded views are learned by the classification model.

For this purpose, a sliding window is employed for sampling regions of same size in both training and testing. The window slides in concentric circles around the center, covering both aperture and particle regions and allowing overlapping. The sliding window is represented in Fig. 5.9.

There is the possibility of regions that cover halfway the aperture and it is not possible to assign a positive or negative label. In such cases, the region is not considered for training in order to avoid noisy fuzzy data. Nevertheless, these regions are allowed to be evaluated by the classifier in the testing phase, whose classification depends on their classification confidence level.

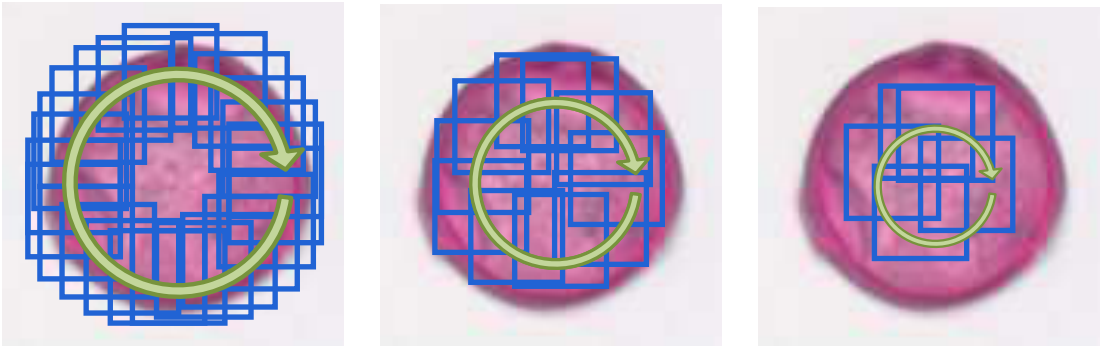


Figure 5.9: The green arrows indicate the movement of the sliding window around the particle in three concentric paths for the sampling process. Overlapping sampled regions are indicated by blue rectangles.

5.7/ Detection of Apertures from an Unknown Image

At this point, the methodology for training the aperture descriptor has been reported. The present section shows how to apply the created codebook and the SVM region classifiers in order to detect apertures on unknown pollen. The evaluation of an unknown particle requires sampling the image in overlapping regions as explained before. Although the method does not require the information about the segmentation of the pollen, it is convenient to employ the contour of the shape to bound the sampling area of the sliding window and hence, to reduce the number of tested regions. The center position of the sampled regions is restrained to the area inside the contour of the pollen estimated in Sec. 4.2.

For each individual classifier the following detection method is applied. First, the aperture descriptor from Sec. 5.4 is computed on each region based on the codebook of visual words. The corresponding SVM classification model is applied to the region descriptor in order to assign a class confidence value. This value can be interpreted as the likelihood of the region to belong to the aperture class. The likelihood is converted to a binary decision by applying a decision threshold (τ_d) so that the regions are labeled either as positive (1) or negative (0). The binary label of each evaluated region is considered as a vote supporting the corresponding class of the region. Because different regions cover the same area due to the overlapping, multiple votes are considered to reinforce the belief of the existence or absence of an aperture in a specific area of the particle.

The method to combine multiple region votes is the creation of a confidence map with the same size as the particle image. Votes of the classified regions are assigned to each pixel of the particle image. The value of the confidence map C for the pixel located at (x, y) is the averaged contribution of all the region votes expressed as

$$C(x,y) = \frac{\sum_{i=1}^{n} V_i}{n}, \tag{5.1}$$

where V_i is the vote of the region i, and n is the total of evaluated regions that contain the pixel (x, y). It is possible to increase the resilience to the misclassification of the regions by increasing the number of evaluated overlapping regions. It can be deducted from eq. 5.1 that the greater n, the less impact of a single region on the final confidence value.

The confidence map can be interpreted as the likelihood of the presence of apertures on the pollen. High confidence areas of the map indicate the location of the detected apertures, which can be found by identifying peaks on the map, *i.e.* local maxima. In order to avoid abrupt changes on the map due to sharp sliding window and to the irregular sampling, the confidence map is smoothed with a Gaussian filter, whose size is one-quarter the size of the regions. Since areas with low confidence values can still show undesired local maxima, the confidence map is trimmed from low values below the confidence threshold (τ_c). Finally, the (x, y) positions of the local maxima are regarded as potential apertures. Figure 5.10 shows examples of the confidence map of different particles and the position of the detected apertures.

From the map confidence map C, two statistical measures are computed to express the content of aperture information on the image. These measures together with the aperture count are considered the set of apeture features to be used in the global classification of pollen taxa.

The first measure is the cumulative sum of all the values of the C. Because each individual classifier contributes with the creation of different maps, the addition over the three maps such that the feature AP_sum is given by:

$$AP\_sum = \sum_{q} \sum_{x} \sum_{y} C_q(x,y), \tag{5.2}$$

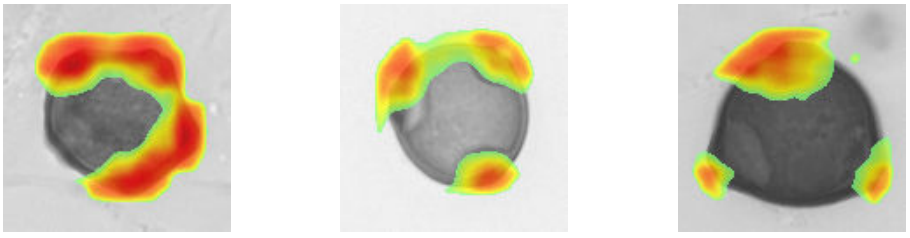
where q is the index for each of the individual classifiers.

The second measure considers the mean value of the confidence map. In order to account for the contribution of each individual classifier, the mean over the three maps C_q is estimated as:

$$AP\_mean = \frac{\sum_{q} \overline{C}_q(x,y)}{Q}, \tag{5.3}$$

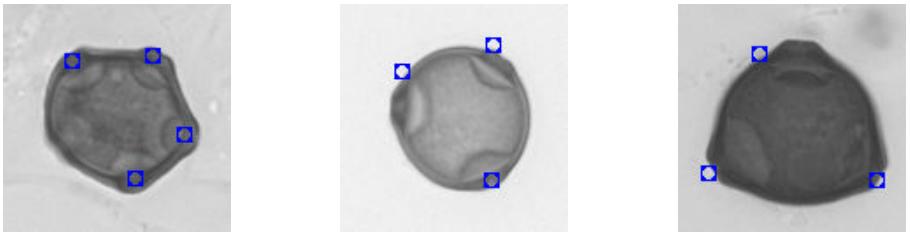
where Q is the number of individual classifiers, $\overline{C}_q(x, y)$ is given by:

$$\overline{C}_q(x,y) = \frac{\sum_{x} \sum_{y} C_q(x,y)}{m}, \tag{5.4}$$



alder birch hazel

(a) Confidence maps



alder birch hazel

(b) Detected apertures

Figure 5.10: Examples of detection of apertures of three pollen types. (a) The confidence maps overlapped over the original pollen image alder (left), birch (center), and hazel (right) pollen. Warm colors (red-orange) indicate high values of the map, regarding a high probability of the presence of an aperture. On the contrary, cold colors (blue-green) indicate low values of the map. (b) Blue marks indicate the position of the correctly detected aperture after finding local maxima on the confidence map of the same particles in (a).

and m is the total number of pixels in C_q.

A final evaluation of the apertures is applied. Apertures too close each other are considered the same in order to reduce multiple detections on the same aperture. This effect could happen when multiple regions around an aperture vote positively generating multiple local maxima. If two or more detections are closer to each other than a given detection distance (δ_d) the detections are regarded as the same aperture, and therefore merged. The measure employs the Euclidean distance. The merging method uses the position of the detections to find a new position that lies in the middle point. The new position is regarded as the single detected aperture and the original apertures are discarded.

Each individual classifier was designed to detect one type of aperture from a single pollen taxon. This means that only one individual classifier is expected to detect apertures on a specific unknown particle, while the rest of classifiers should deliver no results. However, it is possible that two or more classifiers detect different apertures on the same pollen particle due to variability of its appearance. To benefit from these detections, the method gathers the results from all the individual classifiers by performing the union operation on all the individual detected sets, such that the final set of apertures AP_count is formed by all the detections, as depicted in the scheme of Fig. 5.11. This value is considered as the aperture feature for classification purposes.

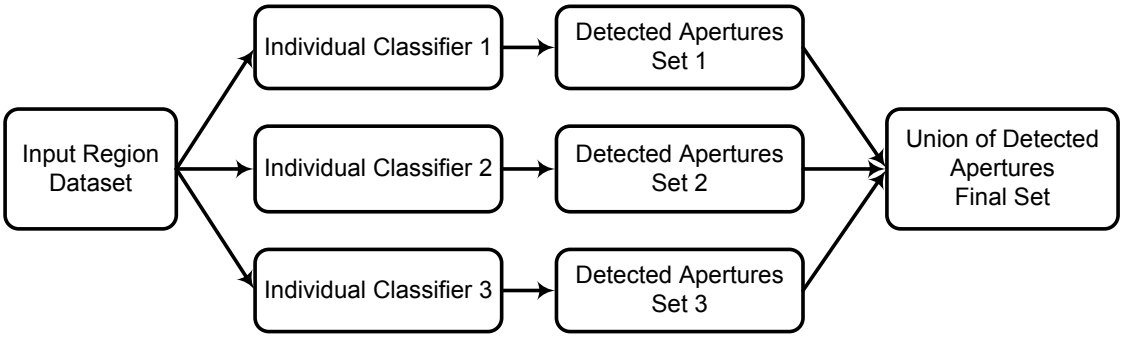


Figure 5.11: Creation of the final detected set of apertures by the union of results of the individual classifiers.

5.8/ Evaluation of the Aperture Detector

5.8.1/ Implementation of the Aperture Detector

The aperture detector is illustrated with a simulation of a real-world application for the evaluation of pollen of the family *Betulaceae*, namely alder, birch, and hazel. In order to evaluate extensively the performance of the aperture detector on real-world applications, additional mugwort and grass taxa are considered in the negative class. Indeed, they have also apertures, but their size, position and visibility are unfavorable. Mugwort apertures are usually hidden in a furrow with no particularly distinctive morphology, therefore difficult to observe. Grass apertures are small compared to the size of the particle, which

makes it hardly visible. The number of apertures available in the dataset is reduced considerably because they show up very rarely. Moreover, the small region required for this type of aperture contains very few visual shape patterns, which limits the robustness of its representation. However, the inclusion of these taxa allows the evaluation of the proposed approach not only in the actual task of detecting apertures, but also in the task of avoiding false detections on other taxa.

Although the aperture features are not capable individually for the recognition of pollen, it is important for distinguishing among taxa that share similar shape and texture.

The evaluation considers the count of apertures AP_count because the comparison to a ground truth of apertures is possible. The other two features, AP_sum, and AP_mean, together with AP_count are incorporated in the global taxa classification in combinations with GSF, EFD, and Texture groups in Ch. 6.

Practical considerations

The source of samples is the single-taxon dataset. A total of 555 particles are employed with the following distribution: alder 115, birch 136, hazel 52, mugwort 120 and grass 132. The total of ground-truth labeled apertures are 541.

Sampling regions from two additional taxa generates an excess of negative samples in contrast to the number of apertures. The raw total of negative sampled regions of the whole dataset is 20368 while the positive samples amount to 2005. This is an important difference and it can be too computational expensive, if all are used for training, especially when multiple parameters are tested, and training cycles need to be repeated several times. For the experiments on the codebook, random sub-sampling is employed to reduce significantly the negative class in order to allow the assessment of multiple configurations with a diminished computational load.

For the parameter adjustment of the SVM classifier, a grid search, in which several parameter combinations are tested within a range of variation, is employed together with a threefold cross-validation. The parameter γ is tested in the range [0.00001 - 1] and C in [0.1 - 1000] with five steps each. A threefold cross-validation means a training-testing ratio of 0.66-0.33. The split of the dataset is stratified, which means that the proportion of positive and negative samples of the original dataset is kept unchanged in the fold's subsets.

5.8.2/ Statistical Performance Measures

This section describes the statistics measures that are employed throughout the experimental part for the configuration and global evaluation of the performance of the detector. The following definitions are based on the work in Ref. [Fawcett, 2003]. Considering the classification of two classes, positive and negative, the following counts are defined:

- False positive (FP) is the count of negative instances incorrectly classified as positive.
- True positive (TP) is the count of positive instances correctly classified as positive.
- False negative (FN) is the count of positive instances incorrectly classified as negative.

- True positive (TN) is the count of negative instances correctly classified as negative.

With these counts, some rates related the total of positive (P) and negative (N) instances can be obtained as follows:

$$\textit{FP rate} = \frac{FP}{N}, \quad \text{and} \quad \textit{TP rate} = \frac{TP}{P}. \tag{5.5}$$

Additionally, the precision combines the TP and FP counts in the following manner:

$$\textit{Precision} = \frac{TP}{TP + FP}. \tag{5.6}$$

Note that the addition of TP and FP stands for the total of detected instances classified as positives.

The F-measure is an additional score that combines the precision and the TP rate. The possible output ranges between 0, in the worst case, and 1, in the best case. The F-measure is given by:

$$\textit{F-measure} = \frac{2}{\frac{1}{\textit{Precision}} + \frac{1}{\textit{TP rate}}}. \tag{5.7}$$

Moreover, a common powerful technique is employed for visualization and comparison of the classification performance with different configurations. In the present study, the Receiver Operating Characteristics (ROC) graph is employed to observe the classification behavior under the changes of threshold parameters. The ROC graph consists of a curve that shows the trade-off between the FP rate on the x-axis and the TP rate on the y-axis.

The ROC curve shows the evolution of the classifier performance (given by the combination of the FP rate and the TP rate) by the variation of the decision threshold to get a discrete classification output. The axes are limited by the possible values of the FP rate, and the TP rate, in the range from 0 to 1. It is possible to plot different ROC curves in the same graph corresponding to different classifiers.

In order to quantize the overall performance of a classifier (for example, for comparison purposes), the Area Under the ROC Curve (AUC) is employed. In the perfect case, the ROC curve wraps completely the graph, which means that the AUC is equal to one. In the worst case, when the classification is random, the ROC curve is a diagonal line cutting in two the graph, with an AUC of 0.5. The AUC can be applied as the single measure of overall performance of a binary classifier, simplifying the comparison.

5.8.3/ Configuration of the Codebook

The first configuration of the aperture detector is the creation of the codebook. In this step, the patch size, the region grid size and the number of visual words of the codebook are chosen based on the performance of the classification of the aperture regions: a region belongs to the positive class if it is an aperture region; otherwise it is considered in the negative class.

Different possible performances are obtained, which correspond to the variation of the threshold that binarizes the SVM classification decision τ_d, and which are plotted in a

ROC graph. The ROC curve represents the global performance of the classifier for all the values of τ_d. Since the detector is tuned at the matching phase, a codebook's classifier that performs best for all τ_d is more significant, independently of τ_c, the threshold employed to determine the significant regions of the confidence map on which local maxima is estimated. The AUC measure is employed as the evaluation measure of the classifier since it summarizes the ROC performance. Therefore, the configuration of the codebook with the best AUC is chosen.

Different region configurations are evaluated with region sizes that enable covering the aperture's area, ranging from 24 to 44 pixels. The codebook size is evaluated in the range of 5 to 25 visual words.

Due to high computational load of the training, the dataset is reduced to 1500 negative samples. The positive class is adjusted to around 500 regions, except for hazel, which does not reach this number (446 samples). In order to avoid bias on a specific subset, regions are randomly sampled. This strategy enables the consideration of the variety of all the regions, especially of the negative set. This decision does not affect the training since the value of the final performance measure is not the most relevant at this stage, but the comparison among region configurations. The average AUC is computed from a threefold cross-validation. Results for each configuration and each individual classifier are shown in Fig. 5.12.

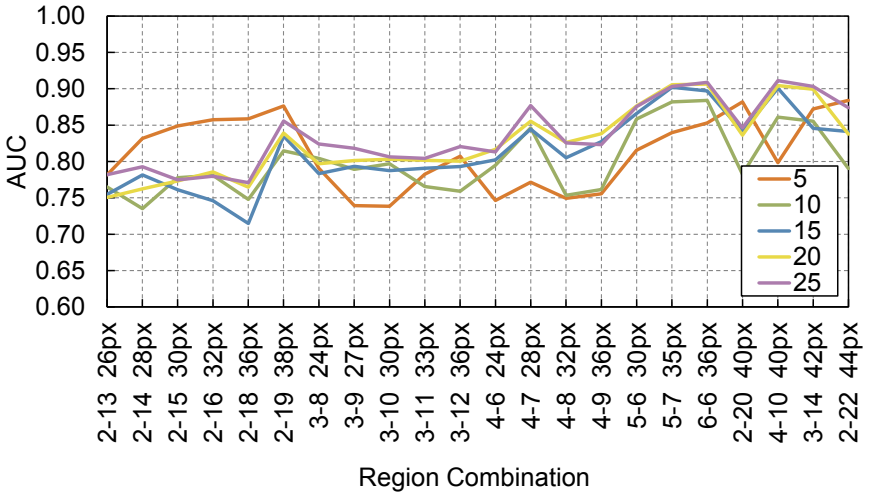
Results show the configurations with best performance. For the alder classifier, the regions with the best performance are configurations 5-7, 6-6, 4-10, and 3-14, which range from 35 to 42 pixels and have an AUC around 0.90. In the case of the birch classifier, the best configurations are 2-19, 4-7, 5-7, 6-6, 4-10, and 3-14 which range from 35 to 42 pixels with an AUC close to 0.95. Finally, for the hazel classifier, the regions 2-19, 6-6, 2-20, and 4-10 show the best results with an AUC around 0.90 in the range of 36 to 40 pixels. These region sizes, from 35 to 38 pixels, are congruent with typical aperture sizes.

It is also noticeable, that bigger codebook sizes have, in most of the cases, better performance that smaller sizes. This is due to the fact that the codebook requires more visual words to represent the variety of the aperture's morphologies. For the aforementioned best configurations, the performance levels become comparable when the codebook size increases, reducing the effect of this selection. A number of visual words $k = 25$ is chosen because it shows the overall best performance for most of the configurations. In all the cases, higher codebook size would mean a more intense computation. The AUC level of the three classifiers for this size is shown in Fig. 5.13.

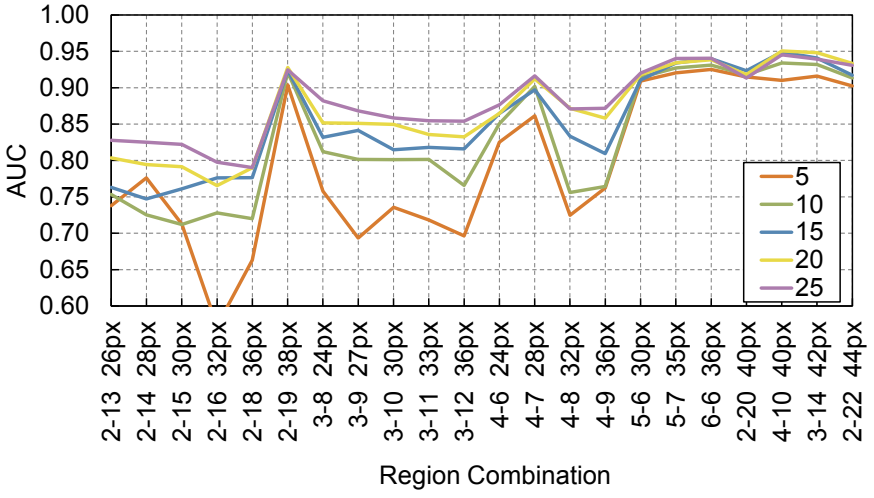
It is convenient to select the same configuration for all the classifiers because the computation load is reduced considerably if the particles are sampled only once with a unique regions size and grid size, instead of sampling three times for each classifier. Under this precept, a single computation of the patch descriptor is useful for the three cases. Therefore, it is advantageous when a single region configuration shows good performance for all the individual classifiers.

Configurations using a grid 2x2 show unstable values when varying the patch and codebook sizes. Although the region 2-19 has particularly good performance, it is discarded since it seems an exception and such as critical selection is risky.

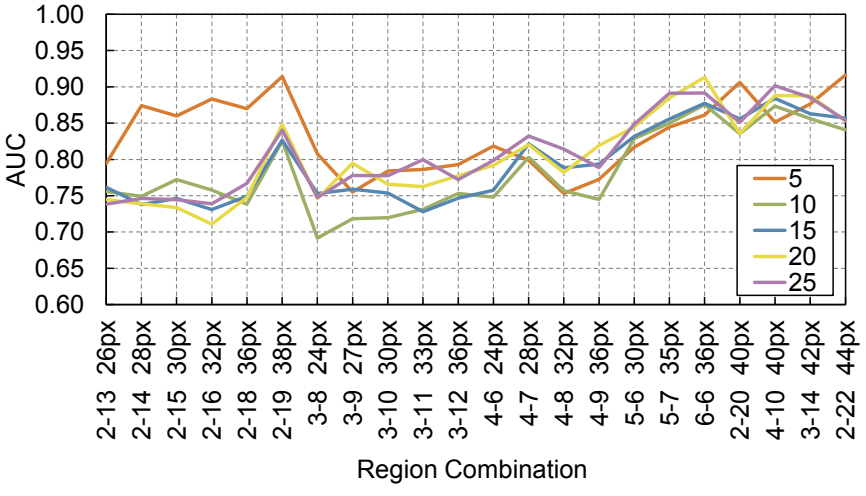
The remaining best regions have a grid size between 4x4 and 6x6. Regions 4-10 and 3-14 are also rejected because the size of the region, bigger than 40 pixels, exceed the size of typical apertures and includes extensive areas belonging to the rest of the particle



(a) alder classifier



(b) birch classifier



(c) hazel classifier

Figure 5.12: AUC performance of the region classification with different size configurations for the individual classifiers: (a) alder, (b) birch, and (c) hazel. Each curve represents the choice of different number of visual words ranging from $k = 5$ to $k = 25$.

and to the background. This reduces the confidence that the regions characterize solely apertures. From the rest of common successful configurations, region 6-6 is chosen due to the high AUC values, although 5-7 could work well too.

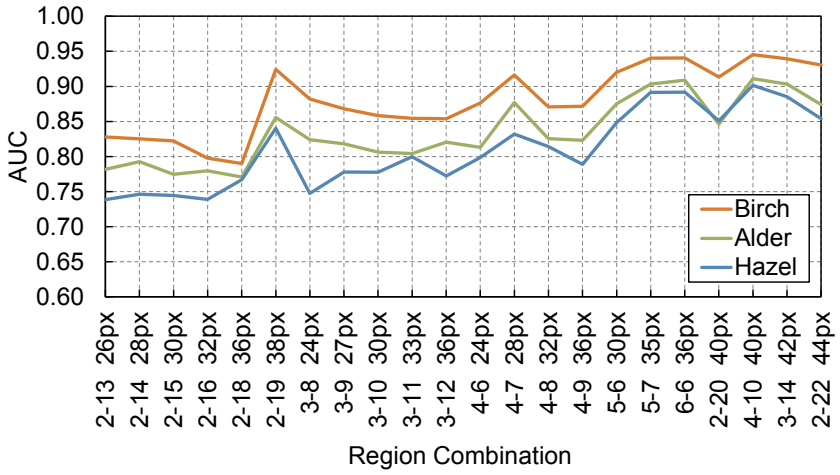


Figure 5.13: A comparative of the three individual classifiers for 25 visual words.

5.8.4/ Strategy for the Evaluation of the Detection

In this section, the procedure of evaluation of the aperture detector is explained based on the confidence map. The first step is to explain some considerations about the reference set of positive regions of the dataset. After that, the rules for scoring the detection are given.

The appearance of the apertures in the image changes gradually depending on the viewing angle. Hence, it is possible to find aperture images from well-defined and sharp to blurred and fuzzy. Additionally, depending on the employed focal plane, the image can offer a partial view of the aperture. While classifiers are designed to recognize only sufficiently-defined apertures, the test of unseen particles employs all the sampled regions without the possibility of rejecting fuzzy or incomplete apertures. For this reason, it is convenient to split the data into two sets according to their visual condition.

Well-defined apertures are employed as the ground truth of apertures for testing, and grouped into the Ground Truth dataSet (GTS). The misdetection of apertures of the GTS is penalized in the evaluation.

Because of the flexibility of individual classifiers to recognize a variety of aperture appearances with visual conditions different from the GTS, it is possible and advantageous to detect some apertures that are partially defined and for which the classifier is not trained. For evaluation of the detector, the apertures that are unexpected to be detected are grouped into the UnExpected aperture dataSet (UES). If the detector misses one of these apertures, there is no penalization. Examples of both sets, GTS and UES, are shown in Fig. 5.14. Note that it is possible to find both cases in the same image.

The evaluation of the performance of the aperture detector consists of quantizing the correct matching of the detected apertures with the evaluation datasets GTS and UES.

alder birch hazel

Figure 5.14: Examples of the aperture sets of alder, birch and hazel used for evaluation. Well-defined apertures marked by a red arrow pointing left belong to the GTS. Fuzzy apertures marked by a blue arrow pointing right belong to the UES.

The performance measure is based only on the matches of the GTS because it is the dataset for which the detector is designed. The estimation of the matches with the UES is carried out in order to allow a correct computation of performance and to gain additional insight into the response of the detector when addressing this kind of aperture.

The rest of this section explains the procedure to match the detected apertures with the reference datasets and how to estimate the different performance measures of the detector. A distance matrix is employed to compare the set of detected apertures (DTS) on one side, to the group defined by the GTS and UES on the other side, for each particle at a time.

First, the closest aperture in the GTS or UES is found for each in the DTS. For this purpose, the Euclidean distance is computed one-to-one between the apertures of both compared sets. An example of this matrix is shown in Fig. 5.15a.

Then, each aperture in the DTS is matched to the closest aperture in the GTS or UES as long as the distance is below an evaluation threshold (δ_{ev}). Note that it is possible to have apertures in the GTS or UES without match in the DTS, as shown in the example of the Fig. 5.15b. Likewise, apertures in the DTS can be unmatched if distances are above δ_{ev}. Finally, statistics of the matches are computed as follows:

- A TP count (*hit*) is regarded for each aperture in the GTS that is matched to at least one aperture in the DTS.
- A FN count (*miss*) is regarded for each aperture in the GTS that is not matched to any aperture in the DTS.
- An unexpected TP count (*unexpected hit*) is regarded for each aperture in the UES that is matched to at least one aperture in the DTS.
- A FP count (*false alarm*) is regarded for each aperture in the DTS that is not matched to any aperture in the GTS or UES.
- A FP count (*false alarm*) is also regarded for each aperture in the GTS or UES that is matched to more than one aperture in the DTS (multiple detection).

The computation of these statistics is exemplified in Fig. 5.15c.

The aperture detector can be seen as a classification task and evaluated with the same performance measures. Then, precision, TP rate, and FP rate are computed from the detector's statistics. True positives of the UES are not considered for the total of true positives, and likewise its false negatives are not penalized since the method is not intended to detect them.

In the next evaluation step, a ROC graph is employed to compare the performance of each classifier for different values of τ_c. In a typical ROC curve, the threshold variation applied to a classification confidence directly affects the trade-off between the TP rate and the FP rate. However, in the case of the aperture detection, the effect of the change of τ_c is quiet more complex because there is an extended process, which involves the search of local maxima, refinement, and matching, before obtaining the DTS. Therefore, the trace of the ROC is not necessarily smooth, especially at the edges where apertures counts can be zero.

	GTA1	GTA2	GTA3	UEA1
DTA1	30	**5**	33	50
DTA2	40	32	**8**	31
DTA3	33	41	**15**	42
DTA4	36	37	44	**12**

(a) distance matrix

	GTA1	GTA2	GTA3	UEA1
DTA1		1		
DTA2			1	
DTA3			1	
DTA4				1

(b) matrix of matches

	Apertures	Total
True Positives	GTA2 GTA3	2
False Negatives	GTA1	1
False Positives	DTA3	1
Unexpected True Positives	DTA4	1

(c) statistics

Figure 5.15: Example of evaluation of four detected apertures (DTA) against three ground-truth (GTA) and an unexpected (UEA) apertures. (a) The distance matrix shows the Euclidean distance between the sets. Minimum distances are highlighted in bold. (b) If the distance is less than δ_{ev} = 27 pixels, DTAs are matched, and marked as 1 in the matrix of matches. (c) Finally, evaluation statistics are extracted from the matrix of matches.

Additionally, the computation of the ROC has to be adjusted in contrast to the typical procedure. The peculiarity is that there is not a limited dataset of the negative class N in 5.5. Actually, N is represented for all the regions of the particle that are not an aperture and it is not fair to use it directly for the computation of the FP rate since it would tend towards zero due to the high number of regions. The solution is to consider the total of apertures in the DTS instead. This adjustment makes the ROC curve more sensitive to the size of DTS, and especially unsteady for a reduced number of detections, which affects the left-bottom side of the ROC.

5.8.5/ Results of the Evaluation

After the evaluation of the codebook, where the region 6-6 and the codebook size of 25 words are selected, the classification performance is evaluated in this section. The robustness of the method to unseen pollen is evaluated by means of a fourfold cross-validation scheme, in which the training-testing ratio is 0.75-0.25. A lower training ratio is not convenient because it would reduce the number positive regions, in addition to the reduction due to the split by the individual classifiers.

For each fold of the cross-validation, the three individual classifiers corresponding to alder, birch, and hazel are created using a SVM classification model with the training data. Their parameters are tuned in a grid search similarly to Sec. 5.8.3.

Due to the split of the training set into the threefold cross-validation of the SVM, a subsampling is no longer considered on the dataset. There are about 15000 negative-class regions for training each fold. Classification confidence values of the testing data are collected after the application of the SVM model. The parameters τ_d and τ_c are left variable in order to compare their effect in a ROC graph.

For the selection of δ_d and δ_c, there is a range in which values are allowed to fluctuate. Due to their intrinsic meaning, which determine the flexibility of the position detected apertures, distances cannot be greater than the typical size of an aperture (38 pixels). Values for δ_d and δ_c are set to 27 and 35, however, further experiments (*cf.* Sec. 5.8.5.1) show the effect of the variation of these parameters on the performance of the aperture detector.

A specific classifier occurrence is given by a fixed τ_d and τ_c and is represented by a point in the ROC curve (FP rate, TP rate). The selection of the optimal classifier can be estimated by the classification costs, which weight independently the importance of the FPs and TPs. In general, the cost function is given by:

$$Cost = \mathit{FPrate} \times c(FP) + (1 - \mathit{TPrate}) \times c(FN), \tag{5.8}$$

where $c(FP)$ and $c(FN)$ are the costs of FP's and FN respectively [Provost and Fawcett, 2001].

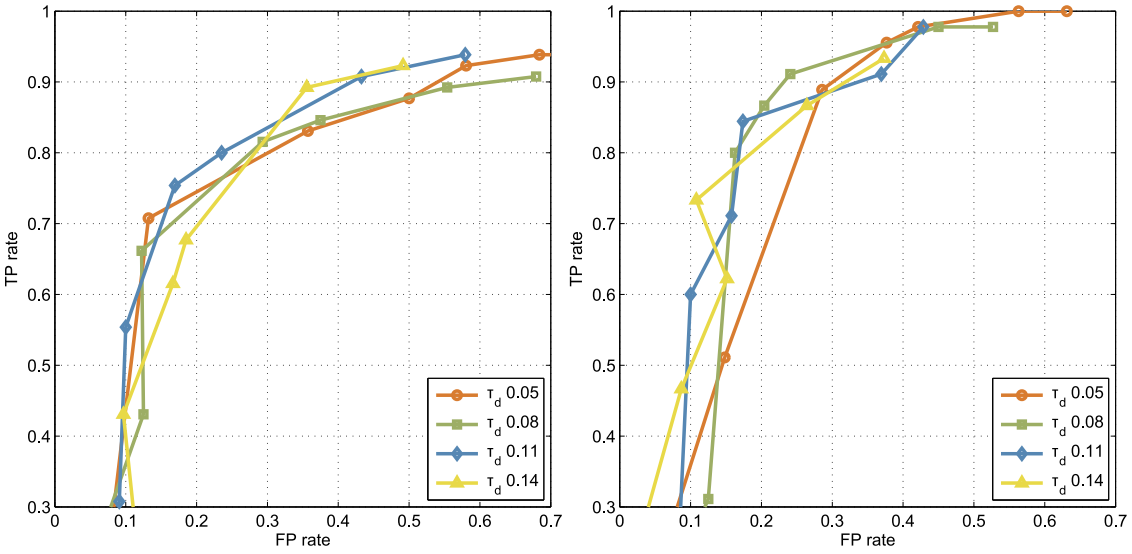
In critical classification tasks, misclassification costs are asymmetrical because they lead to significant actions. For example, in cancer diagnosis or poison detection, the cost of FNs must be much higher than that for FPs. For the case of pollen classification, specifically for aperture detection, misdetections are not critical since the classification decision is supported by additional information (for example, shape and texture). Moreover, the misclassification of a single pollen particle is neither determinant for the concentration rate nor dangerous for the patient's health.

For purposes of this study, costs $c(FP)$ and $c(FN)$ are assumed to be symmetrical. This means that the cost of a missed detection is equal to the cost of a false alarm. With this condition, the optimal detector is the one found at the closest point (FP rate, TP rate) to the upper left corner of the ROC graph in any direction.

Figures 5.16(a)-(c) show the ROC graph corresponding to a fold test. The complete set of results for all the folds is found in the Appx. A. Each ROC curve is computed against the variation of τ_c for each individual classifier. A family of ROC curves is created for different values of τ_d. Values of τ_d are relatively low compared to typical values in a binary classification (around 0.5) due to the unbalance of the dataset. It is also noticeable that in some cases, the selection of τ_d is not too critical and different values have similar results. Performance of the hazel detector is somewhat lower than the rest. This difference can be attributed to the smaller size of the dataset for this class.

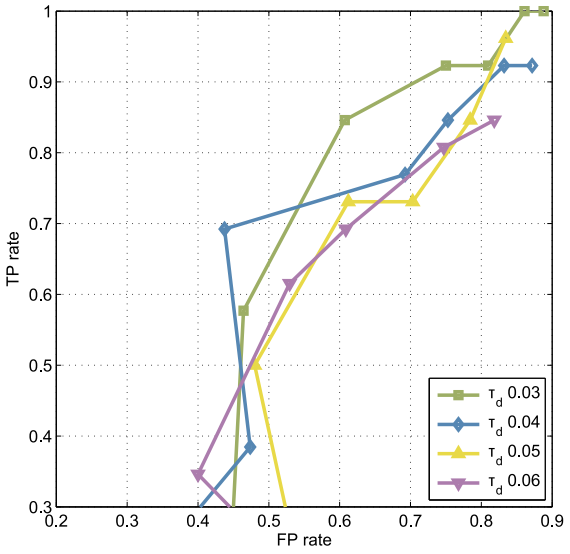
The precision and the F-measure are employed to assess the performance of the overall cross-validation test of the aperture detector. The precision is computed as the average performance from the four folds considering all the individual classifiers, after setting a particular combination of τ_d and τ_c. The F-measure is computed using the averaged statistics.

The following results are given with a confidence level of 0.95. Considering the classification of the five pollen types, average precision is 0.70 $\pm$ 0.06, the average TP rate is 0.78



(a) alder classifier

(b) birch classifier



(c) hazel classifier

Figure 5.16: ROC curves of evaluation on the five-taxon aperture detection for the three individual classifiers for one of the fourth folds: (a) alder, (b) birch, and (c)hazel. Variation of τ_c creates different points of the curve. Variation of τ_d is represented by a set of different ROCs.

± 0.03, the FP rate is 0.29 ± 0.06. This performance achieves an F-measure of 0.74. Limiting the calculation to only the three learned pollen types, joint performance shows an average precision of 0.76 ± 0.05, TP rate remains 0.78 ± 0.03, but the FP rate moves to 0.23 ± 0.05, while the F-measure reaches 0.77. The difference between both results is caused by FPs within mugwort and grass sets, which do not contribute with TP counts. This effect can be explained by the flexibility allowed during the training to compensate the high variability of the aperture appearances. This flexibility causes that some artifacts on the morphology of mugwort and grass particles are misidentified as apertures. Higher resolution images describing more information of apertures could help in these difficult tasks. Also recall that the adjusted computation of the FP rate is more sensitive to misclassifications.

5.8.5.1/ EFFECT OF δ_d AND δ_{ev}

In this section, the effect of the variation of δ_d and δ_{ev} on the performance of the aperture detector is evaluated. The configuration of each classifier is chosen to have the same values of τ_d and τ_c that in Sec. 5.8.5. Although the natural range of δ_d and δ_{ev} is bounded to the typical size of the apertures (38 pixels), the test covers up to 41 pixels. The tables in Appx. B show the results for each combination in the range from 25 to 41 pixels.

High values of δ_{ev} allow more flexibility to the detector regarding the position of the detected aperture and increase the TP rate. However, this flexibility also increases the FP rate. In contrast, high values of δ_d decreases the TP rate while increasing the FP rate. The precision shows a similar behavior to the FP rate. It is important to show that the variability of the statistics, which is measure by the ranges of the confidence intervals, tends to reduce for the best performances levels of each case.

Finally, the F-measure combines the behavior of the precision and the TP rate. Different maximum values are found in the region corresponding to high δ_{ev}, and medium to high δ_d.

Although configurations beyond 38 pixels tend to have a better performance, it is not wise to employ them because the distance error between the positions of the detected aperture and the actual apertures is allowed to be greater than the size of the aperture itself.

5.9/ CONCLUSIONS

Important elements are introduced to the method in order to improve its robustness. The use of a BOW codebook enables the detector to represent apertures with a variety of appearances. Moreover, the split into individual classifiers is important for optimizing the detection by aperture type, and for adding models for new types.

The experiments confirmed that the aperture detector is suitable for overcoming the intra-class variation and inter-class similarity to conditions comparable to real-world applications and even with added extra pollen taxa with no visible apertures. Moreover, the cross-validation evaluation showed that the method is robust to unseen particles. The method succeeded under total lack of previous knowledge about location and aperture quantity and under challenging conditions such as multiple input taxa and analysis on a

single focal plane.

Some flexibility is allowed for the configuration of the codebook, enabling to choose a single region size for all the classifiers. Similarly, the settings of the distance thresholds can be adjusted to the expected aperture size for optimal results. The parameters τ_d and τ_c are selected such that a minimal classification cost is given. The cost function can be modified for example, based on the allergenicity and frequency of the taxa, following discretion of allergologists and palynologists.

It was also demonstrated that the training requires a great number of samples due the high variation of morphological appearances. In the case of the hazel classifier, the detector had a decrease in performance for this reason.

The proposed method can be enhanced by adding detection capabilities at different focal planes, since apertures change their appearance when observed at these planes. One path is the confirmation of apertures already detected and the detection of previously disregarded apertures by independent detections on multiple planes. A second path is the volumetric description of the apertures using the information from different focal planes, for example the application of a volume LBP operator as the local descriptor.

6

Relevance Analysis of Features and Global Classification

Continuing with the last stage of the recognition system described in Ch. 4, the analysis of the extracted features and the global classification of pollen are conducted in this chapter. The purpose is to identify the features and the setup that better account for the differences among pollen taxa, and hence, are more relevant for the discrimination and classification of the studied allergenic pollen.

The purpose of the analysis is to filter out those features that contribute little or nothing to the recognition of the most important pollen taxa in terms of allergenicity. The analysis enables also the understanding of the major characteristics (shape, texture, apertures, and size) that describe the pollen.

The strategy of analysis is the application of feature selection methods that test many different combinations of features within the characteristic groups. In particular, the family of *wrapper* methods is employed for this purpose [Kohavi and John, 1997]. The characteristic of this family is that the relevance of the features is measured based on their impact on the classification result.

Wrappers repeat a cycle in which they create and evaluate a classifier model with a specific subset of features. In each cycle, the subset is modified (by adding or subtracting features) and performance is compared to the previous cycle. The process is stopped until little or no improvement is achieved with the new subset. Therefore, the remaining feature subset is considered the optimal subset for that classification task. After the selection of relevant features for each characteristic group, the combination of groups is evaluated in a global classification scheme.

The rest of the chapter is organized as follows. First, the feature selection methods are introduced. Then, the characteristic groups are analyzed and relevant features are determined. Later, experiments on the global classification at taxon level are conducted, involving the relevant features. Finally, conclusions from the classification of pollen are drawn.

6.1/ Feature Selection Methods

Wrappers are applied to the analysis of pollen features because they account for the interaction among features based on the idea of classification of pollen taxa as a joint set,

which is the main purpose.

Three feature selection methods are employed. First, Brute Force (BF) consists of the most exhaustive search of features. The method evaluates all possible combinations of features until finding the subset that outputs the best performance. Due to the number of combinations, the method is only affordable for a reduced number of features.

For the case of big sets of features, Sequential Forward Selection (SFS) and Recursive Feature Elimination (RFE) are employed. SFS method begins with a small subset and grows it as long as the performance increases. In contrast, RFE works backwards by reducing the whole feature set. The main advantage of both cases over BF is that not all possible combinations are evaluated.

6.1.1/ Brute Force

This is clearly the optimal solution to maximize the performance, since all possible combinations of features are tested. Given the feature set $X = \{x_1,\ x_2,\ \ldots,\ x_m\}$, BF starts searching in the smallest subset $\{x_k\}$ for any feature k, testing all m features individually and keeping the results. Then, all the combinations $\{x_{k1},\ x_{k2}\}$ of any two features are created and tested. The cycle continues until the whole set X containing all the features is tested. Finally, all the performances are compared to pick the subset with the best performance. In this approach, there is no stopping criterion.

The only condition that can be used to avoid testing all the combinations is the limitation of the maximum number of features. With this consideration, BF tests only subsets as greater as the defined parameter.

Because the computational load is high, BF can be used only for a small number of features. For this reason, the method is employed only for the analysis of General Shape Features (GSF).

6.1.2/ Sequential Forward Selection

The concept behind SFS is to start from the smallest optimal subset of features which is grown one feature each iteration such that the performance is maximized, and to stop until a better performance is not possible [Sergios Theodoridis, 2006]. It starts similarly to BF, testing all sets $\{x_k\}$. SFS determines the best one-feature subset $\{x_{w1}\}$ based on the performance. Then, only subsets consisting of the combination $\{x_{w1},\ x_k\}$ of the previous best subset and any other feature k are evaluated. The winning combination $\{x_{w1},\ x_{w2}\}$ with best performance is kept as the new best subset. The cycle is repeated increasing the size of the best subset by one feature each iteration until there is not any new subset that provides better performance that the current selection.

As in BF, an additional stop criterion can be considered by setting the maximum number of features that the winning subset can contain.

The computational load of SFS is reduced compared to BF and it is suitable for the evaluation of high dimensional feature sets. One drawback of the SFS is that it does not always find the best subset, because it is dependent on the selection of features on the first rounds. However, it provides most of the time a very good estimation of the optimal subset.

6.1.3/ Recursive Feature Elimination

RFE is another wrapper approach that works in opposite direction to SFS: it starts from the whole set X and removes irrelevant features until having the best performance [Guyon et al., 2002]. In order to avoid a high load of computation for high dimensional sets, the method removes multiple features in each iteration instead of only one.

A second important difference is the criterion employed for determining which features are eliminated. RFE employs a SVM classification of the evaluated subset to rank the features individually. Specifically, the weight magnitude of the SVM is employed as ranking criterion since it expresses the effect on the objective function of removing features. RFE takes out the features with lowest ranking, leaving a more relevant subset $\{x_{w1}, \ldots, x_{wk}\}$. Typically, 50% of the features are dropped in each iteration. As in previous selection methods, the maximum number of selected features can be chosen to force an actual reduction of features and to find the most relevant subset.

6.2/ Analysis of Characteristic Feature Groups

The analysis of the features is conducted in groups organized in the same manner that they are described in Ch. 4. This division enables a clearer understanding of the descriptive power of the major characteristics as well as the specific feature measures that are relevant for discrimination.

The application of the aforementioned feature selection methods, according to the dimensionality of the groups, enables the assessment of features in each group. In a final batch of experiments, the relevant features that are found in this step are employed together on the global classification of pollen taxa.

The strategy of analysis is as follows. The performance criterion for the selection methods is the accuracy of the classification using a SVM model with a RBF kernel function. First, the parameters of a SVM classifier are optimized using a grid search as described in Sec. 5.1. The parameter γ is tested in the range [0.1 - 1000] and C in [1 - 15000] with 20 steps each. For this purpose, all the features of the evaluated group are input.

In the next step, the corresponding feature selection method is applied using again SVM classification. By using the already computed optimal SVM parameters, the effective selection of features is not affected since they are compared under the same classification model while avoiding the computation of optimal parameters in each cycle.

Finally, the performance of the selected subset of features of the group is evaluated by SVM classification, with SVM parameter grid optimization. In order to avoid overfitting, in all the cases when SVM classification is employed, the cross validation scheme is applied with five folds. This strategy implies that for each fold, a partition of size four fifths of the master set is randomly sampled and assigned to the training set. The remaining fifth is employed as the testing set. The first condition for sampling is that a particle cannot be part of more that one testing set. This means partitioning the original set in complementary subsets, and then combining them to form the training and testing sets for each fold. The second condition is that the training and testing sets must preserve the same rate of particles of each pollen type as the original set.

A total of 448 particles are employed from the single-taxon dataset with the following

distribution: alder 100, birch 100, hazel 48, mugwort 100 and grass 100. The particles were randomly selected from the original 555 of the single-taxon dataset in order to have an even distribution, except for hazel, for which 100 samples were not possible.

6.2.1/ General Shape Features

The group of general shape features described in Sec. 4.4.1, which is listed in 4.1 and 4.2, is evaluated using the BF feature selection method due to its low dimensionality.

Some features are removed from this group because of their dependence on the size of the particle, namely *perimeter, area*, and *2eN* features. The purpose of this action is to obtain a group that bases the description of the pollen shape solely on morphological characteristics, that is, only on the shape of the particle and there is no bias related to differences in size among taxa.

Having the final evaluated group made of eleven features, the BF is computationally affordable. Results are evaluated at different values for the maximum allowable number of features in the BF method, which changes in consequence the subset size of selected features. The method is likely to get a smaller set of features than the maximum if the classification performance is better. Results of this experiment are shown in Table 6.1. The performance with all eleven features (without selection) is also shown for comparison.

Table 6.1: Classification performance of the GSF group using the BF method with different maximum allowable features. The actual number of employed features is indicated in each case.

Max. feat.*	3	5	8	11	No selection
Act. feat.**	3	5	6	9	11
Accuracy	0.806 ± 0.016	0.839 ±0.047	0.857 ± 0.031	0.857 ± 0.023	0.841 ± 0.041
R				•	•
U		•		•	•
ratio1		•		•	•
ratio2	•		•	•	•
ratio3	•	•	•	•	•
rdis	•	•	•	•	•
cxf			•	•	•
EF_rms			•	•	•
EF_mean		•	•		•
EF_std				•	•
EF_ratio					•

* Maximum allowable features.
** Actual number of features.
• Selected feature.

Some conclusions can be drawn from these results. Even the smaller selected subset of three features is able to achieve a good classification accuracy of 0.806. Moreover, with the addition of some more features, accuracy gets to 0.857. This means that the most important information is contained in these three features, which can be considered the basic set of strongly relevant features: *ratio2, ratio3*, and *rdis*. This subset is present in

almost all other configurations. The rest of the features contain complementing information that helps increasing the performance, which can be considered as weakly relevant features.

Compared to the classification with no selection, some subsets can achieve slightly better results, meaning that an appropriate selection of features is important.

Comparing the subsets of five and six selected features, performance is very similar, although with a quite different subset in which only three features are common. This result suggests that some features are equivalent and they could be exchangeable in terms of classification.

In contrast, the subsets of three and six features are much more similar. In the latter set, the basic set of three features is kept with the addition of three features (*cxf, EF_rms*, and *EF_mean*), which enable to increase the performance by 0.051.

In order to discover if there is a feature in particular responsible of that increment, the basic set of three features is tested individually with each of the additional features from the subset of six features. Results in Table 6.2 show that features related to the error of the fitted ellipse are the major individual contributors of the improvement. *EF_std* is also evaluated in a similar manner because of its close relation to these features. A comparable result indicates that *EF_std* has a similar effect to *EF_rms* and *EF_mean* and it could be exchangeable as well.

The performance of the four-feature subsets is slightly superior to the one obtained by BF with five and six features. This result does not indicate that BF has failed. Recall that the given accuracy is computed in a second cross-validation round for optimizing SVM parameters, which could result into differences of the performances if different parameters are chosen. However, the SVM parameter configuration is always the same in the local BF process to maintain comparison fairness.

Table 6.2: Classification performance of the combination of the basic subset of three features with each of the Ellipse-fitting features.

Number of features	4	4	4	4
Accuracy	0.808 ± 0.019	0.859 ± 0.027	0.850 ± 0.024	0.857 ± 0.022
R				
U				
ratio1				
ratio2	•	•	•	•
ratio3	•	•	•	•
rdis	•	•	•	•
cxf	•			
EF_rms		•		
EF_mean			•	
EF_std				•
EF_ratio				

• Selected feature.

The subset with the four following features shows to be effective and compact: *ratio2, ratio3, rdis*, and *EF_rms*. Hence, it is chosen to represent the general shape characteristic of the pollen in further experiments in the global classification.

6.2.2/ EFD Features

The analysis of the EFD features is conducted with a different strategy due to the high dimensionality of the group. SFS and RFE methods are employed since the evaluation of all the feature combinations is impracticable. Employing the first 150 EFD features described in Sec. 4.4.2, SFS and RFE selection methods are applied independently. In both cases, the performance is evaluated against the variation of the maximum number of selected features. As in BF, this parameter makes possible a smaller set of features compared to the initial maximum.

Performance of SFS on the EFD group is shown in Table 6.3. Comparison with the performance with all 150 features (without selection) is added to the table.

Table 6.3: Classification performance of the EFD group using the SFS method with different maximum allowable features. The actual number of employed features is indicated in each case.

Maximum allowable features	Actual selected features	Accuracy
10	10	0.813 ± 0.022
50	44	0.846 ± 0.044
100	44	0.846 ± 0.044
150	44	0.846 ± 0.044
No selection	150	0.772 ± 0.036

The performance progressively increases starting with a small subset of ten features up to an optimal point where SFS does not select more features. This best subset consists of 44 features. The best performance is not achieved using all the EFD's. On the contrary, the whole set makes the performance to decrease. A basic subset of ten strongly relevant features is able to achieve an accuracy of 0.813. The remaining 34 features contribute only with a performance increment of 0.033.

A very similar behavior is observed when using the RFE method as shown in Table 6.4. Here, the basic subset of ten EFD's leads to an accuracy of 0.812, while the best performance is achieved with 50 features (six more than SFS) for an accuracy of 0.852. This means that the contribution of the 40 extra features yields to the improvement of the performance by 0.040. The classification with 44 selected features is also computed to allow comparison with results of SFS. Here, RFE is slightly lower.

Table 6.4: Classification performance of the EFD group using the RFE method with different maximum allowable features. The actual number of employed features is indicated in each case.

Maximum allowable features	Actual selected features	Accuracy
10	10	0.812 ± 0.047
44	44	0.839 ± 0.042
50	50	0.852 ± 0.048
100	100	0.819 ± 0.033
No selection	150	0.772 ± 0.036

The comparison between the two methods is illustrated in Fig. 6.1. Both methods behave

practically in the same way, even though the selected subsets are not identical.

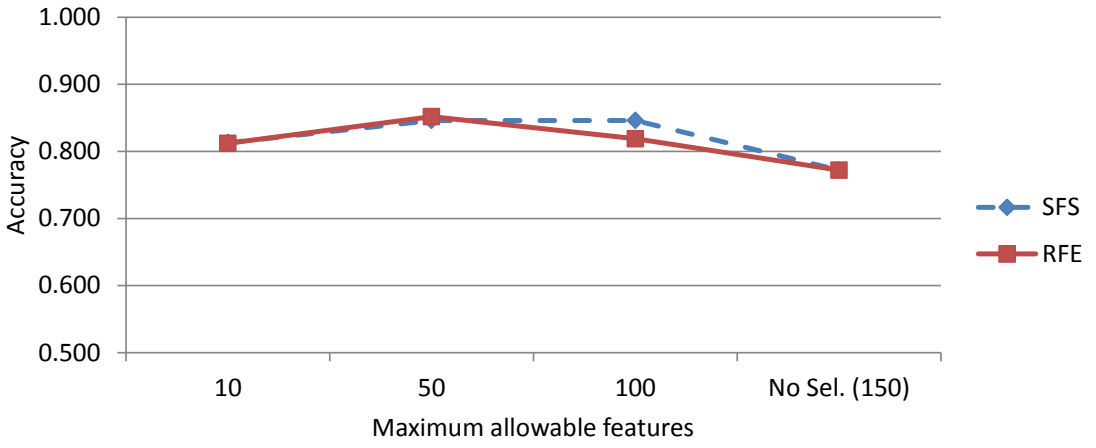


Figure 6.1: Performance comparative between SFS and RFE selection methods at different choices of the maximum allowable features for the EFD group.

The actual selected features in both cases are similar, and both methods tend to select the first harmonics more relevant than the last ones. The basic subset of ten features in both methods is selected from the first 13 harmonics. For comparison purposes, the best subsets of each method are analyzed: 44 features for SFS and 50 for RFE. In Table 6.5, the relative histogram of the selected features according to their harmonic number is indicated for the best subsets of SFS and RFE.

Table 6.5: Relative histogram of the EFD's selected by SFS and RFE methods. The distribution is based on the EFD number of the descriptor.

Bin	SFS	RFE
1st-38th harmonic	36%	44%
39th-75th harmonic	16%	16%
76th-113th harmonic	32%	22%
114th-150th harmonic	16%	18%

In both methods, the first bin, which represents the first 38 harmonics, contains the most of the EFD's, which matches with the finding of a basic relevant subset. This subset contains the low-frequency information of the contour and outlines its shape.

Interestingly, the second most important bin is the third one, which describes higher frequencies of the contour. This subset contains finer information of the particle, enabling more sophisticated representation of the particle details.

Despite the finer information, the differentiation among taxa provided by this subset is not much superior compared to the case of GSF. This result suggests that the intra-class variability, especially in the high-frequency harmonics, is high enough to prevent an outstanding classification. Therefore, the profit from the fine EFD representation is moderate.

For the global classification, the RFE subset of 50 features is chosen to represent the

EFD group due to its high performance and reduced number of features.

6.2.3/ Texture Features

The analysis of the texture features follows a similar strategy to the EFD case due to the high dimensionality of the group.

According to Sec. 4.4.3, 275 features are extracted consisting of eleven Haralick's measures for 25 different offsets. Thus, SFS and RFE selection methods are applied independently and the variation of the maximum number of selected features is tested. As in previous cases, a smaller set of features is possible compared to the initial maximum.

Results of the application of the SFS method on the texture group are indicated in Table 6.6. Performance with all 275 features (without selection) is compared in the same table.

Table 6.6: Classification performance of the Texture group using the SFS method with different maximum allowable features. The actual number of employed features is indicated in each case.

Maximum allowable features	Actual number of features	Accuracy
10	10	0.766 ± 0.013
55	53	0.797 ± 0.041
110	91	0.815 ± 0.041
165	91	0.815 ± 0.041
No selection	275	0.824 ± 0.028

Table 6.6 reveals a gradual increase of the performance from the subset of ten features up to some optimal subset, smaller than the whole set without selection. However, it is remarkable that even with the ten-feature subset an accuracy of 0.766, indicating the possibility of a basic subset consisting of few strongly relevant features.

The rest of the features, reduce their importance by contributing less to the performance. For example, the following 42 features improve the performance only by 0.031. This effect continues until reaching the 91 selected features with an improvement of 0.049. After this subset size, SFS does not select more features although a higher maximum is chosen, as in the case of allowing 165 features. Finally, the maximum classification performance occurs when 275 features are employed with no selection.

In order to have an accuracy above 0.800, the subset of 91 features is selected the best configuration for SFS. This selection reduces considerably the original feature set and reduces only the performance by 0.009 compared to the maximum with 275 features.

The second evaluation of the feature selection, shown in Tab. 6.7, is conducted with the RFE method. An additional test is conducted with 91 features in order to compare to the same configuration in SFS. Similarly to SFS, the addition of features improves the performances, reaching the maximum when all features are employed.

The overall result of RFE is slightly superior to SFS, as represented in Fig. 6.2. The actual subsets selected by the methods are not the same, although they share similar characteristics. Table 6.8 compares the Haralick's measures from the 91 selected features in both methods. The frequency of the measures is indicated for each method. All Har-

Table 6.7: Classification performance of the Texture group using the RFE method with different maximum allowable features. The actual number of employed features is indicated in each case.

Maximum allowable features	Actual number of features	Accuracy
10	10	0.788 ± 0.046
55	55	0.815 ± 0.041
91	91	0.821 ± 0.028
110	110	0.826 ± 0.034
165	165	0.819 ± 0.004
No selection	275	0.824 ± 0.028

alick's measures show to be relevant at different level. While RFE considers *difference variance, contrast, sum entropy, difference entropy*, and *inverse difference moment* the most important features, SFS considers *inverse difference moment, correlation, entropy, angular second moment, difference variance, sum entropy*, and *sum of squares* in this order.

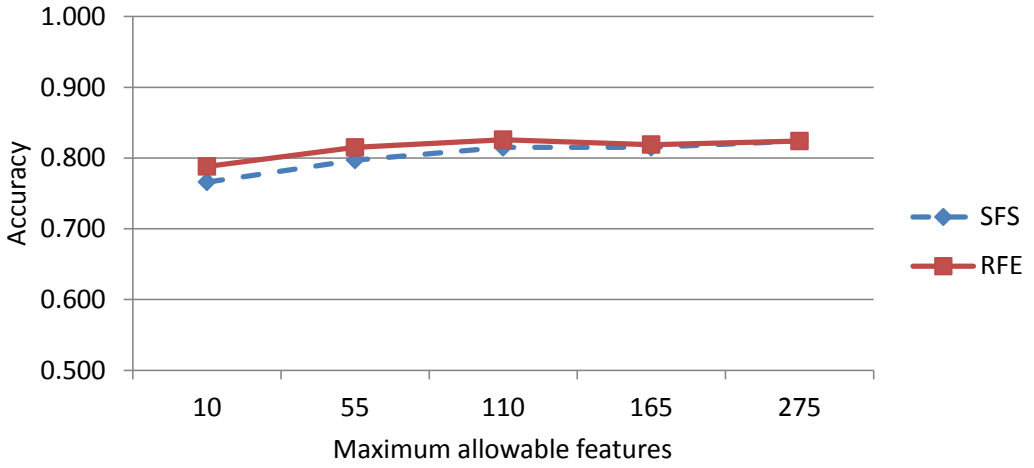


Figure 6.2: Performance comparative between SFS and RFE selection methods at different choices of the maximum allowable features for the Texture group.

A similar comparison is performed focusing on the offsets. Table 6.9 indicates the frequency of selection by each method. From 25 available offsets, SFS considers all, while RFE employs 23. The offset frequency is quite different between the two methods. For RFE, the most important offsets are {5 px , 45°}, {1 px , -90°}, {5 px , -45°}, {2 px , -45°}, and {1 px , 90°}, while for SFS are {3 px , -45°}, {4 px , 45°}, and {4 px , 0°}.

The application of practically all the types of offsets and Haralick's measures reflects the complexity of the pollen patterns. The difference between the selected offsets and Haralick's measures in both experiments, keeping a similar performance, suggests that information of the patterns is redundant and therefore, the features are exchangeable. Hence, it is possible to create different relevant sets with different texture features as done by SFS and RFE methods.

The exact number of selected features for the representation of the texture group is not

Table 6.8: Frequency of Haralick's measures of the 91 features selected by SFS and RFE methods. The maximum frequency is 25, which is given by each of the offsets.

Haralick's measure	SFS	RFE
Difference Variance	**9**	**18**
Contrast	7	**17**
Sum Entropy	**9**	**14**
Difference Entropy	4	**13**
Inverse Difference Moment	**12**	**12**
Angular Second Moment	**10**	4
Correlation	**11**	4
Entropy	**10**	4
Sum Average	5	2
Sum of Squares (variance)	**9**	2
Sum Variance	5	1

Table 6.9: Frequency of offsets of the 91 features selected by SFS and RFE methods. The maximum frequency is eleven, which is given by each of the Haralick's measures.

Offset	SFS	RFE
3 px , -45°	**7**	3
4 px , 45°	**6**	5
4 px , 0°	**6**	1
1 px , 0°	5	5
2 px , 0°	5	5
2 px , -90°	5	5
4 px , -45°	5	2
3 px , 45°	5	1
5 px , 45°	4	**8**
1 px , -90°	4	**7**
5 px , -45°	4	**6**
3 px , 90°	4	4
1 px , 45°	4	4
2 px , 45°	4	3
4 px , -90°	4	1
1 px , 90°	3	**6**
2 px , -45°	3	**6**
1 px , -45°	3	5
2 px , 90°	2	4
3 px , -90°	2	3
5 px , -90°	2	3
3 px , 0°	1	2
5 px , 0°	1	2
5 px , 90°	1	0
4 px , 90°	1	0

critical and depends on the trade-off between of accuracy and size of the subset. Moreover, results exhibit that subsets of configurations with fewer selected features are also part of bigger subsets of configurations with more selected features. Thus, strongly relevant features are re-utilized and complemented with other weakly relevant features in bigger configurations. This fact indicates consistency of the importance of the strongly relevant features.

For the global classification described in the next section, the subset of 91 features obtained by the RFE method is applied as texture representation.

6.3/ Global Classification

The second stage of the analysis is the joint evaluation of the features on the global classification of pollen taxa. The purpose of this test is to observe how different pollen characteristics interact. Due to the different descriptive power of the features, it is expected that the joint contribution provides better classification performance than their individual application.

All the characteristic groups are considered in different combinations. In particular, the combination of the chosen feature subsets of the groups with the apertures features is evaluated first. After the groups GSF, EFD, and Texture are employed jointly, the features related to the apertures and to the size are added progressively to the classification. Finally, the whole set of 151 features from all the characteristic groups, which are summarized in Table 6.10, is evaluated.

Table 6.10: Summary of the five types of features employed in the global classification. A total of 151 features are considered.

Group	Features
GSF	ratio2, ratio3, rdis, and EF_rms
EFD	50 selected features
Texture	91 selected features
Aperture	AP_count, AP_sum, and AP_mean
Size-related	perimeter, area, and 2eN

6.3.1/ Aperture Features with Characteristic Groups

The first experiments consist of the evaluation of the novel aperture features in combination with other groups. The same SVM-cross-validation scheme of evaluation described in Sec. 6.2 is followed. The evaluation employs the accuracy as the performance measure, although average of precision, recall, and F-score are given as additional measures.

Aperture features are coupled with each main group to evaluate their contribution to recognition. The performance of each combination is shown in Table 6.11. In all the cases, results indicate that the aperture features contributed to increase the accuracy of the classification in comparison to the classification by individual groups. This comparative is depicted in Fig 6.3. The contribution of the aperture features to the accuracy is in

average 0.094.

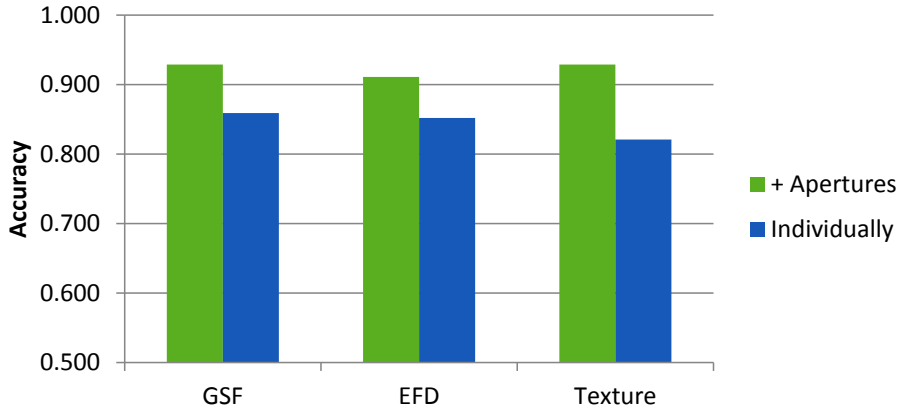


Figure 6.3: Contribution of the aperture features to the classification performance. The accuracy of the combination of the feature groups and the aperture features is contrasted to the individual accuracy of the groups. Aperture features prove to enhance the individual classification models.

Table 6.11: Classification performance of the combination of the apertures features with the rest of characteristic groups.

Accuracy	0.929 ± 0.020	0.911 ± 0.040	0.929 ± 0.009
Precision (avg.)	0.929 ± 0.025	0.919 ± 0.027	0.917 ± 0.059
Recall (avg.)	0.910 ± 0.075	0.898 ± 0.054	0.912 ± 0.071
F-score (avg.)	0.917 ± 0.044	0.906 ± 0.021	0.915 ± 0.065
GSF	•		
EFD		•	
Texture			•
Aperture	•	•	•

The confusion matrices of the three cases are shown in Table 6.12. The distribution of the errors is quite even among the different taxa in GSF and EFD cases, except for nine hazel particles classified as alder in Table 6.12a and nine alder particles classified as mugwort in Table 6.12b. This result could suggest that morphological variations of alder are strong enough to be similar to other pollen morphologies.

Table 6.12c shows that most of the misclassifications fall among the genera of the family *Betulaceae*, while mugwort and grass and almost perfectly classified. This result confirms the similarity of the texture patterns inside the same family and the difference compared to the rest.

The increase of the performance in the three cases clearly evidences the relevance of the Aperture group. Moreover, the content of information carried by their features complements the pollen description based on typical shape and texture.

Table 6.12: Classification confusion matrices of the individual combination of the characteristic groups with the Aperture group.

		True				
		Alder	Birch	Hazel	Mugwort	Grass
Predicted	Alder	**93**	3	9	0	1
	Birch	3	**91**	0	0	2
	Hazel	2	1	**36**	0	0
	Mugwort	2	2	2	**100**	1
	Grass	0	3	1	0	**96**

(a) GSF + Aperture

		True				
		Alder	Birch	Hazel	Mugwort	Grass
Predicted	Alder	**89**	5	3	2	0
	Birch	2	**90**	5	0	1
	Hazel	0	1	**38**	0	0
	Mugwort	9	2	0	**93**	1
	Grass	0	2	2	5	**98**

(b) EFD + Aperture

		True				
		Alder	Birch	Hazel	Mugwort	Grass
Predicted	Alder	**90**	5	6	0	1
	Birch	5	**91**	5	0	1
	Hazel	5	3	**37**	0	0
	Mugwort	0	0	0	**100**	0
	Grass	0	1	0	0	**98**

(c) Texture + Aperture

6.3.2/ All the Characteristic Groups

The second and final experiment is the combination of all the groups in order to seek the maximum performance. Starting with groups GSF, EFD and Texture as a base, groups Aperture and Size-related are added individually. Finally, the combination of all the five groups together is evaluated. The results are shown in Table 6.13.

Table 6.13: Performance of the combination of all the characteristic groups in the classification of the five pollen taxa.

Accuracy	0.958 ± 0.022	0.969 ± 0.021	0.969 ± 0.025	0.982 ± 0.006
Precision (avg.)	0.960 ± 0.029	0.968 ± 0.020	0.970 ± 0.028	0.980 ± 0.016
Recall (avg.)	0.960 ± 0.029	0.968 ± 0.022	0.968 ± 0.025	0.982 ± 0.012
F-score (avg.)	0.960 ± 0.029	0.968 ± 0.020	0.969 ± 0.022	0.981 ± 0.012
GSF	•	•	•	•
EFD	•	•	•	•
Texture	•	•	•	•
Aperture		•		•
Size-related			•	•

The solely combination of the groups GSF, EFD and Texture achieves an excellent accuracy of 0.958 with 19 misclassifications. Moreover, the consideration of additional information, given by the Aperture and Size-related groups, seems crucial to maximize the classification capabilities of the model. The addition of the group Aperture or the Size-related contributes with 0.011 to the accuracy. Finally, the classification with all the five groups yields to the best performance with an accuracy of 0.982 and with eight misclassifications. The confusion matrix for these configurations is shown in Table 6.14.

From Table 6.14a, it is evident that most of the misclassifications are between alder and birch (14 out of 19). The problem is mitigated when the groups Aperture or Size-related are considered (nine and eleven out of 14). Moreover, in the most complete configuration, which considers the addition of both groups at the same time, the number of particles in this situation is only five out of eight.

For this configuration, seven out of eight cases fall among genera of the family *Betulaceae*. These facts are in line with the presumption that genera of the same family are mutually similar and more difficult to classify. These experiments confirm that the consideration of additional descriptive groups is important for the discrimination of difficult cases among similar taxa.

Focusing on the results of the most complete configuration illustrated in Table 6.14d. Fig. 6.4 presents the eight particles that are misclassified. The only pollen of this set that is not part of the family *Betulaceae* is one mugwort particle. This particle is very distinctive among the mugwort pollen, with an atypical black area that affects more drastically the texture pattern; fact that explains the classification failure.

The misclassification of the *Betulaceae* particles can be explained by the inter-class similarity, with the help of a comparative among particles of alder, birch and hazel. By inspecting row-wise Fig. 6.5, the inter-class similarity of genera of this family is patent. In the misclassified cases, the method cannot simply adjust the boundaries among classes because it needs to compensate the intra-class variability of each genus, which is evident when comparing column-wise.

Table 6.14: Confusion matrices of the classification of the five pollen taxa using different combinations of the five characteristic groups together.

		True				
		Alder	Birch	Hazel	Mugwort	Grass
Predicted	Alder	**92**	7	1	0	1
	Birch	7	**92**	0	0	0
	Hazel	0	0	**47**	1	0
	Mugwort	1	0	0	**99**	0
	Grass	0	1	0	0	**99**

(a) GSF, EFD, Texture

		True				
		Alder	Birch	Hazel	Mugwort	Grass
Predicted	Alder	**96**	5	1	0	0
	Birch	4	**93**	1	0	0
	Hazel	0	1	**46**	1	0
	Mugwort	0	0	0	**99**	0
	Grass	0	1	0	0	**100**

(b) GSF, EFD, Texture + Aperture

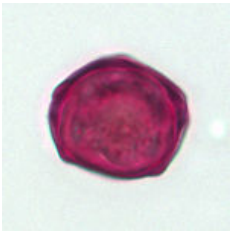
		True				
		Alder	Birch	Hazel	Mugwort	Grass
Predicted	Alder	**92**	3	1	0	0
	Birch	8	**97**	1	0	0
	Hazel	0	0	**46**	1	0
	Mugwort	0	0	0	**99**	0
	Grass	0	0	0	0	**100**

(c) GSF, EFD, Texture + Size-related

		True				
		Alder	Birch	Hazel	Mugwort	Grass
Predicted	Alder	**98**	3	1	0	0
	Birch	2	**96**	0	0	0
	Hazel	0	1	**47**	1	0
	Mugwort	0	0	0	**99**	0
	Grass	0	0	0	0	**100**

(d) GSF, EFD, Texture + Size-Related + Aperture

Alder classified as birch



Alder classified as birch

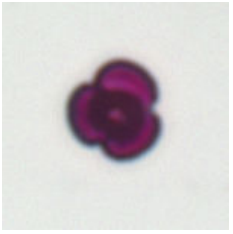
Birch classified as alder

Birch classified as alder

Birch classified as alder

Birch classified as hazel

Hazel classified as alder



Mugwort classified as hazel

Figure 6.4: The eight misclassifications from the global model of Table 6.14d.

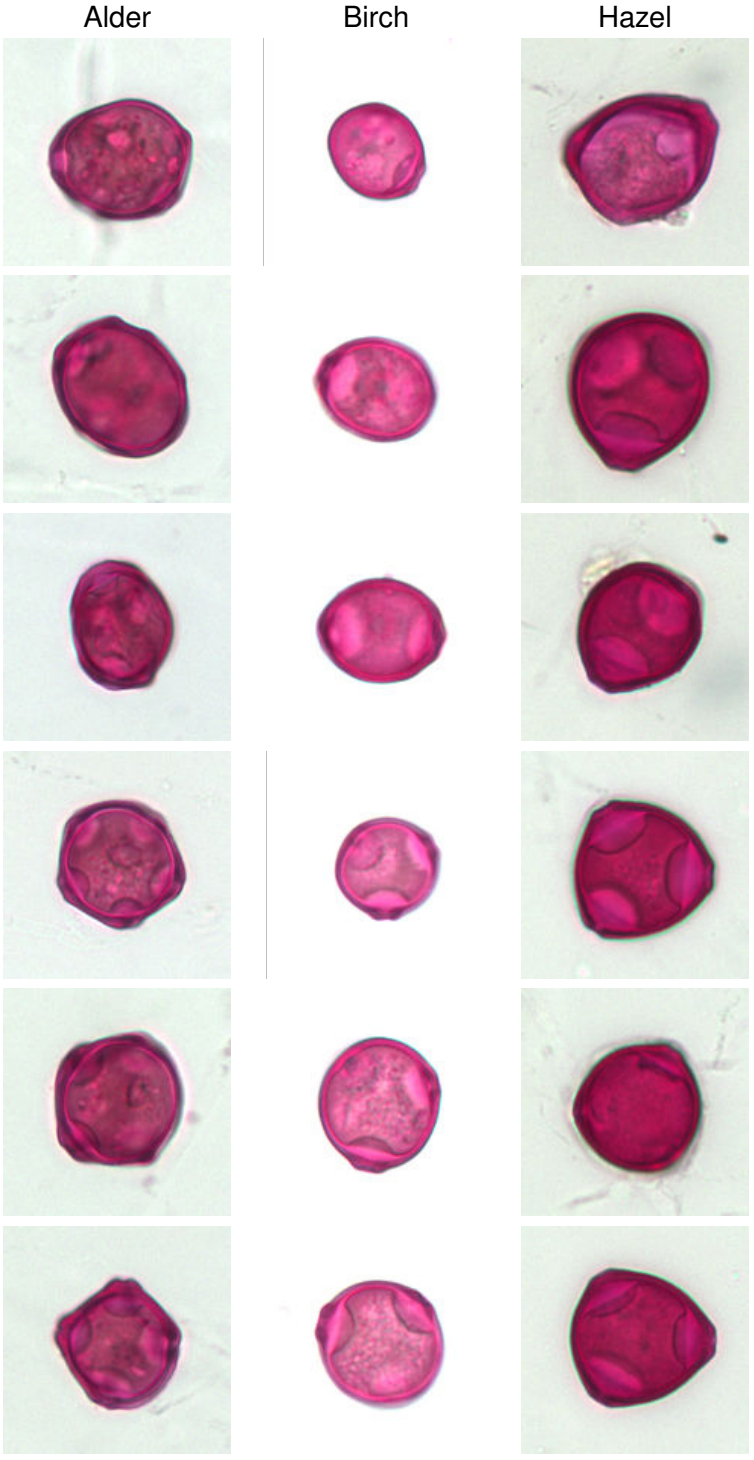


Figure 6.5: Comparative of the inter-class similarity and intra-class variability of the family *Betulaceae*. Examples of genera alder, birch, and hazel are shown in each column.

6.4/ CONCLUSIONS

In this chapter, the comprehensive evaluation of features with focus on the correct description of pollen particles, according to their characteristic groups, was conducted. Relevant subsets of each group were recognized in the first batch of experiments and finally they confirmed suitability for working together.

Results report that each descriptive group provides critical information for the recognition of pollen taxa. With the appropriate selection of relevant features, the set of features can be reduced and the classification performance maintained or even improved.

Moreover, characteristic groups complement mutually deficiencies by improving their discrimination capacity when combined. In particular, the aperture features presented this effect when combined with other characteristic groups, due to their capability of provide additional information that is rarely considered.

It is clear that the most robust condition is achieved when the diversity of the groups work together, achieving excellent performance. It was observed that the discrimination power of the method is affected mostly by the similarity of the genera of the same family, which is manifested as misclassified particles among the three taxa that correspond to the family *Betulaceae*. However, the inter-class similarity was successfully overcome thanks to the contribution of the multiple characteristic groups.

7

CONCLUSIONS

7.1/ RESULTS

The thesis presented an approach for recognition of allergenic pollen taxa from microscope images. It was proposed a set of algorithms for localization and segmentation of pollen particles in airborne slides based on the common characteristics of most of the pollen. A collection of features corresponding to the discriminative characteristics of the taxa were sought with the purpose of thoroughly describing the pollen, which include shape (GSF and EFD), texture, and size.

An exhaustive analysis led to the discovery of strong relevant features through the application of feature selection methods to characteristic groups independently. Subsets of different size were compared to evaluate the impact of features on the classification performance. Results led to the definition of strongly relevant features that were common to most of the subsets and which are indispensable for the recognition of taxa. Moreover, a second subset of features showed weak relevance. These features are needed to improve the classification performance, although less significantly than the strongly relevant features.

To complement the characteristic groups, a novel strategy was designed, evaluated and applied with the goal of detecting different pollen apertures. Aperture features were designed, including the explicit count of the apertures on the particle.

A method for training and creation of the correct model was proposed. The correct parameters were tuned by an iterative process based on the classification performance. A good recognition of apertures was confirmed despite the diversity of other similar patterns formed in the pollen image.

The aperture features were proven to be suitable to boost taxon discrimination on global tests when applied together other feature groups. The aperture detector has the ability to be adapted for the robust recognition of different types of apertures just by applying the training data without modification of the underlying method.

The method was tested on the recognition of the most allergenic pollen taxa in Germany, whose importance is shared in the rest of Europe and the World, specifically alder, birch, hazel, mugwort and grass. The application of the most relevant features from the characteristics groups permitted excellent classification results in combination with proved machine learning algorithms like SVM. Overfitting was avoided by the employment of the cross-validation scheme. The conjunction of diverse characteristic groups was important to improve the performance in comparison to the individual application of the groups.

The proposed approached is among the best pollen recognition processes, with an accuracy of 98.2%. Recalling results from Sec.2.2.8 , the proposal is comparable to the accuracy of 97.2% by Chen *et al.*, with the difference that this study consider two more taxa, a bigger dataset and the use of a single method for aperture detection [Chen et al., 2006]. Similar results were also achieved by Zhang *et al.* with an accuracy of 97.7% and by Ronneberger *et al.* with a precision of 98.5% and recall of 86.5% [Zhang et al., 2004; Ronneberger et al., 2007]. The main difference is that the latter employed 33 pollen taxa in multiple-layer approach.

The performance shown in this thesis are above the those obtained by France *et al.*, Ranzato *et al.*, and Rodríguez Damián *et al.* [France et al., 1997; Ranzato et al., 2007; Rodríguez-Damián et al., 2003].

The proposed strategy proved to be robust against the complexity of task:

- Intra-class variation, considering a natural random position of the pollen particle with respect to the viewing point which affects the pollen morphology as well as the visibility and appearance of the apertures. This is affected also by the natural variability present in biological individuals that belongs to the same genus.
- Inter-class similarity due to the close relation of three of the five pollen genera, which belong to the same taxonomic family.
- Important rate of other particles and debris with respect to pollen in the airborne samples.
- Although the datasets were obtained under similar conditions, there is not a formal standard method for dyeing and digitizing yet. Therefore the processing of samples at early stages is prone to error.

In contrast, the localization and segmentation method struggles within a clutter of pollen, where stuck particles can be considered as a single bigger object and rejected due to its size or shape.

The recognition system is not explicitly prepared to handle correctly occlusion or stuck debris on the pollen particles. In a first filter, the localization method rejects the particles that are too different from the expected types. If the affected particle can still reach the classification stage, the numeric features could change sufficiently to misclassify the taxon. However, the diversity of the characteristics groups works favorably in this direction, by overcoming the deviation of the affected features from their characteristic values. Moreover, because this error is expected to occur randomly on any pollen taxon, the relative pollen rate is not affected by the misclassification of these cases.

The aperture detector can still improve its performance. The consideration of a bigger and more diverse training set may alleviate this situation by providing even more information about the variety of appearances. Nevertheless, the importance of the aperture features was proved on the global classification, showing that the current performance level is sufficient for a profitable contribution.

It would have been desirable to evaluate the whole chain of classification from the extraction of airborne pollen from real samples to the estimation of the final pollen type concentration. Unfortunately, reliable data with expert ground-truth labeling was not possible to gather. This evaluation would allow the performance comparison to the current estimation by palynologists and other systems.

7.2/ FUTURE WORK

The characterization of the pollen can be strengthened by the incorporation of multiple image layers at different focal planes during the image acquisition. This approach would provide richer information of the particle, enabling the confirmation of classification decision through the evaluation of features from different layer.

Moreover, the aperture detector could validate detected apertures and even discover new apertures with more views or the same particle. An alternative improvement is the substitution of the current local feature by a 3D or multiple-layer descriptor, which would enable a volumetric analysis.

This multiple-layer extension requires the evaluation of the quality of the layers, mainly regarding sharpness. The cost of the improvement is an increase in the digitization time, and the growth of the processing.

The incorporation of the sampling date to the set of features is an important consideration. The information can be compared to the pollination calendars of the pollen taxa in order to determine their probability in the slide. Hence, it would enable the discrimination of particles with similar appearance but with different pollination periods resembling palynologists' work.

The definition of a standard for the dyeing process and for the setup of the image acquisition would mean a great step for extending the use of the method, which would enable the analysis of allergenic pollen taxa and their regional variations. The incorporation of additional pollen information would allow for the increase of the robustness and recognition capabilities of the system.

7.3/ PERSPECTIVES

With focus on the development of a fully automatic pollen recognition system, the creation of a smooth process pipeline of the training of the classification models would be favorable. A friendly user interface would enable to palynologist to use the recognition system without further assistance by simply inputting the pollen data and labels. In a similar way, the aperture detector should be automatically modeled from the correct labeled aperture regions.

Regarding to the practical utilization of the system, the implementation of such smooth pipeline would allow for the immediate application of the trained models on the taxon recognition in real environments.

Alternative classification approaches could be considered by taking advantage of the analysis of features in relation to the characteristic groups described in this study. Modern knowledge-based recognition of objects are able to link knowledge from different domains to numeric processing [Truong, 2013]. For example, expert knowledge in palynology may be related to the pollen properties through semantic meaningful concepts in a structured framework. Pollen properties can be defined based on the relevant numeric features analyzed in this study. Thus, a classification based on expert knowledge and a reasoning process would employ the relationships among the pollen concepts.

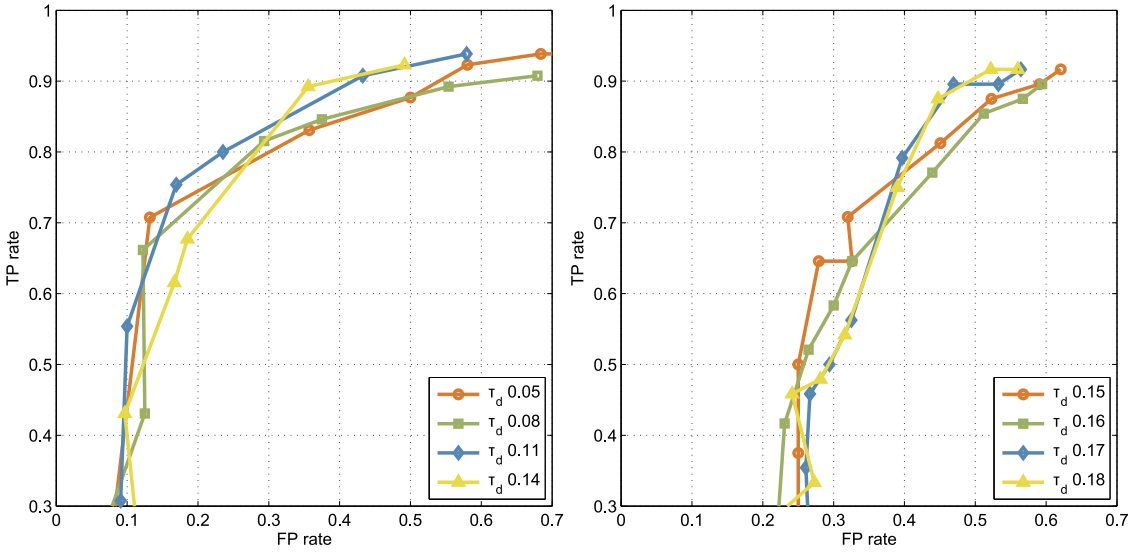
Finally, the strategy applied in this study, and especially the aperture detector, can be ex-

trapolated to other microscopic objects straightforward, for example in the case of spores and cells, which morphology can be described by characteristic groups similar to those presented in this thesis.

A

Performance of the Individual Classifiers

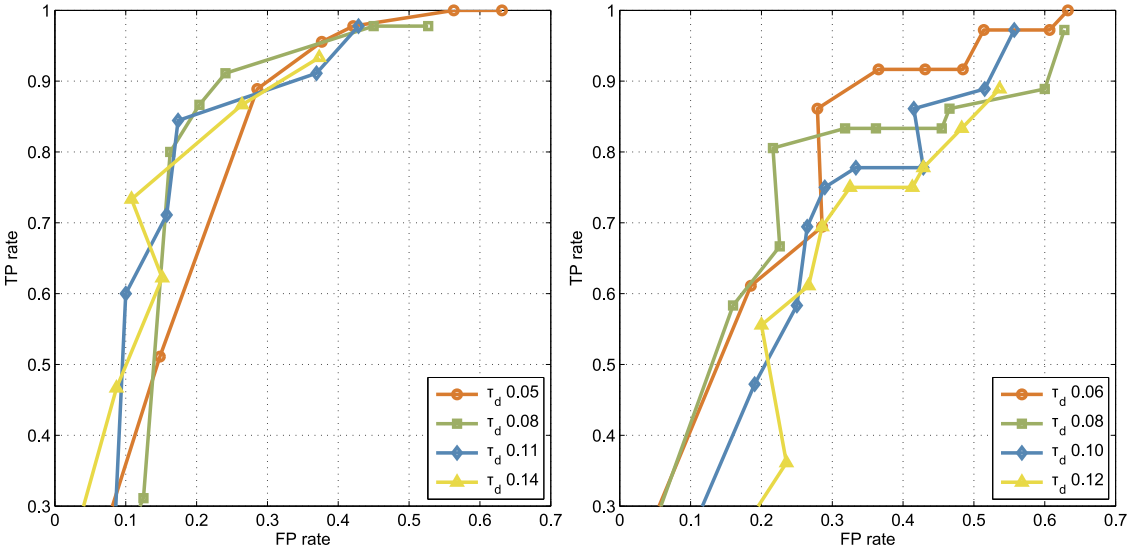
The complete results of the individual classifiers designed in Sec. 5.8.5 and evaluated in five pollen taxa are given in this appendix. For each of the three classifiers (alder, birch, and hazel), ROC curves are created by the variation of τ_c. Different curves are shown at different levels of τ_d. Graphs for each fold are given in all the cases.



(a) Fold 1 (b) Fold 2

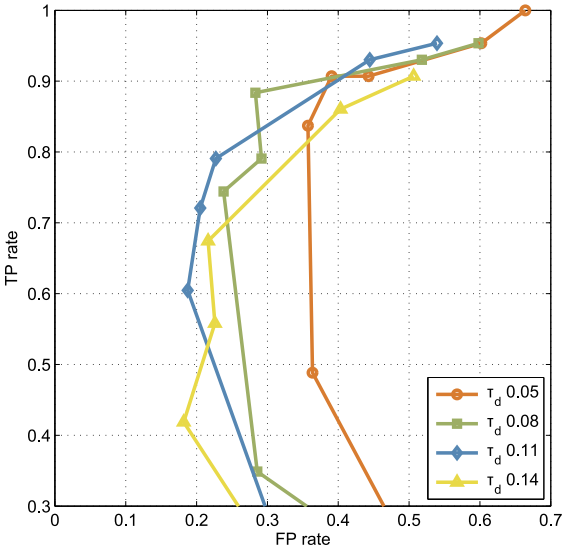
(c) Fold 3 (c) Fold 4

Figure A.1: ROC curves of evaluation of the alder individual classifier.

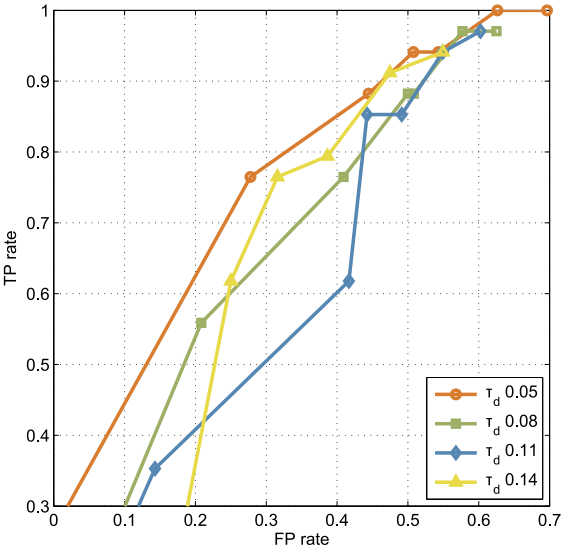


(a) Fold 1

(b) Fold 2

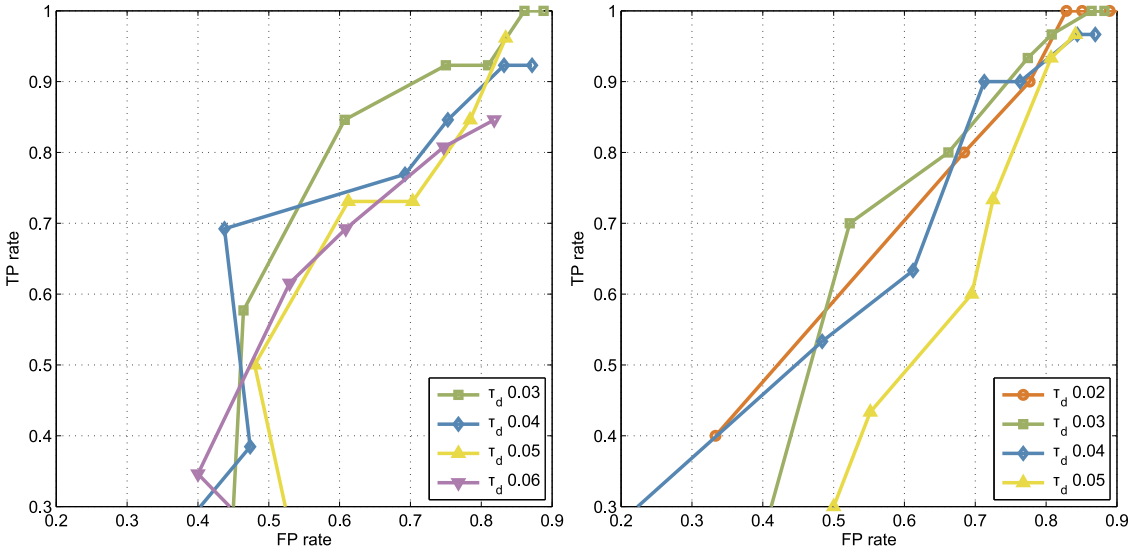


(c) Fold 3



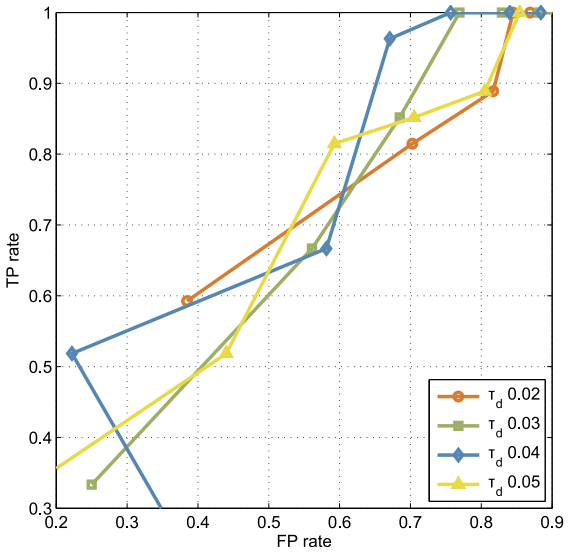
(c) Fold 4

Figure A.2: ROC curves of evaluation of the birch individual classifier.

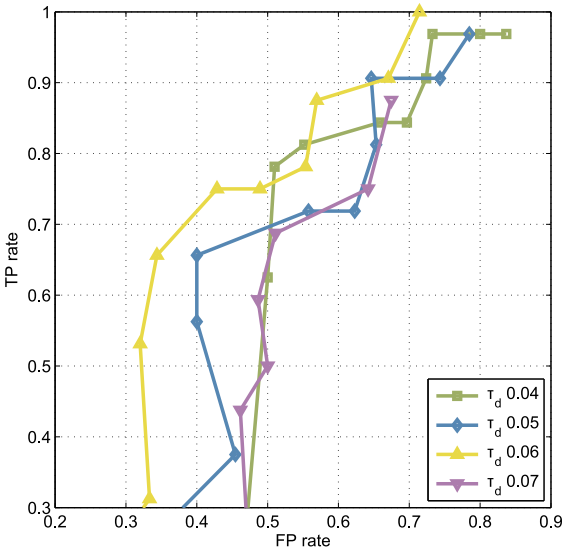


(a) Fold 1

(b) Fold 2



(c) Fold 3



(c) Fold 4

Figure A.3: ROC curves of evaluation of the hazel individual classifier.

B

EFFECT OF δ_d AND δ_{ev}

Tables B.1-B.4 show the effect of the variation of δ_d and δ_{ev} on the performance of the aperture detector in different combinations in the range of 25 to 41 pixels described in Sec. 5.8.5.1.

Table B.1: (a) Variation of the mean TP rate for all the folds in relation to changes in δ_d and δ_{ev}. (b) Respective ranges (±) of the confidence limits of the mean TP rate. In both tables, levels with higher performance are shaded in green while those with lower one are shaded in red.

δ_{ev} \ δ_d	25	27	29	31	33	35	37	39	41
25	0.72	0.71	0.71	0.70	0.69	0.69	0.68	0.68	0.67
27	0.75	0.74	0.73	0.73	0.72	0.72	0.71	0.70	0.70
29	0.77	0.76	0.75	0.75	0.74	0.74	0.73	0.72	0.72
31	0.78	0.78	0.77	0.76	0.75	0.75	0.74	0.73	0.73
33	0.78	0.78	0.77	0.77	0.76	0.75	0.74	0.74	0.73
35	0.79	0.78	0.77	0.77	0.76	0.76	0.75	0.74	0.74
37	0.79	0.78	0.78	0.77	0.77	0.76	0.76	0.75	0.74
39	0.79	0.79	0.78	0.78	0.77	0.76	0.76	0.75	0.75
41	0.79	0.79	0.78	0.78	0.77	0.76	0.76	0.75	0.75

(a) Mean TP rate

δ_{ev} \ δ_d	25	27	29	31	33	35	37	39	41
25	0.05	0.05	0.05	0.06	0.07	0.07	0.07	0.06	0.07
27	0.03	0.03	0.04	0.04	0.05	0.05	0.05	0.05	0.05
29	0.03	0.02	0.03	0.04	0.05	0.05	0.05	0.04	0.04
31	0.03	0.02	0.03	0.04	0.05	0.05	0.04	0.04	0.05
33	0.03	0.03	0.03	0.04	0.05	0.05	0.05	0.04	0.05
35	0.03	0.03	0.03	0.04	0.05	0.05	0.05	0.04	0.05
37	0.03	0.03	0.03	0.04	0.05	0.05	0.05	0.04	0.05
39	0.03	0.03	0.04	0.04	0.05	0.05	0.05	0.04	0.05
41	0.03	0.03	0.04	0.04	0.05	0.05	0.05	0.04	0.05

(b) Confidence intervals

Table B.2: (a) Variation of the mean FP rate for all the folds in relation to changes in δ_d and δ_{ev}. (b) Respective ranges (±) of the confidence limits of the mean FP rate. In both tables, levels with higher performance are shaded in green while those with lower one are shaded in red.

δ_{ev} \ δ_d	25	27	29	31	33	35	37	39	41
25	0.34	0.32	0.31	0.30	0.31	0.30	0.30	0.29	0.29
27	0.31	0.29	0.28	0.27	0.27	0.27	0.26	0.26	0.26
29	0.28	0.27	0.26	0.25	0.25	0.24	0.24	0.23	0.23
31	0.27	0.25	0.24	0.22	0.23	0.22	0.22	0.21	0.21
33	0.26	0.24	0.23	0.22	0.22	0.22	0.21	0.21	0.21
35	0.25	0.23	0.22	0.21	0.21	0.21	0.20	0.20	0.20
37	0.24	0.22	0.21	0.20	0.20	0.19	0.19	0.18	0.18
39	0.24	0.22	0.21	0.20	0.19	0.19	0.18	0.18	0.17
41	0.24	0.22	0.21	0.19	0.19	0.19	0.18	0.18	0.17

(a) Mean FP rate

δ_{ev} \ δ_d	25	27	29	31	33	35	37	39	41
25	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04
27	0.06	0.06	0.06	0.05	0.05	0.05	0.04	0.05	0.04
29	0.06	0.06	0.06	0.06	0.06	0.05	0.05	0.05	0.04
31	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.04	0.03
33	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.04	0.03
35	0.04	0.05	0.05	0.04	0.04	0.04	0.04	0.04	0.03
37	0.04	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.03
39	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03
41	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.04	0.03

(b) Confidence intervals

Table B.3: (a) Variation of the mean precision for all the folds in relation to changes in δ_d and δ_{ev}. (b) Respective ranges (±) of the confidence limits of the mean precision. In both tables, levels with higher performance are shaded in green while those with lower one are shaded in red.

δ_{ev} \ δ_d	25	27	29	31	33	35	37	39	41
25	0.66	0.68	0.69	0.70	0.69	0.70	0.70	0.71	0.71
27	0.69	0.71	0.72	0.73	0.73	0.73	0.74	0.74	0.74
29	0.72	0.73	0.74	0.75	0.75	0.76	0.76	0.77	0.77
31	0.73	0.75	0.76	0.78	0.77	0.78	0.78	0.79	0.79
33	0.74	0.76	0.77	0.78	0.78	0.78	0.79	0.79	0.79
35	0.75	0.77	0.78	0.79	0.79	0.79	0.80	0.80	0.80
37	0.76	0.78	0.79	0.80	0.80	0.81	0.81	0.82	0.82
39	0.76	0.78	0.79	0.80	0.81	0.81	0.82	0.82	0.83
41	0.76	0.78	0.79	0.81	0.81	0.81	0.82	0.82	0.83

(a) Mean precision

δ_{ev} \ δ_d	25	27	29	31	33	35	37	39	41
25	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04
27	0.06	0.06	0.06	0.05	0.05	0.05	0.04	0.05	0.04
29	0.06	0.06	0.06	0.06	0.06	0.05	0.05	0.05	0.04
31	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.04	0.03
33	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.04	0.03
35	0.04	0.05	0.05	0.04	0.04	0.04	0.04	0.04	0.03
37	0.04	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.03
39	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03
41	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.04	0.03

(b) Confidence intervals

Table B.4: (a) Variation of the F-measure for all the folds in relation to changes in δ_d and δ_{ev}. Levels with higher performance are shaded in green while those with lower one are shaded in red.

δ_{ev} \ δ_d	25	27	29	31	33	35	37	39	41
25	0.69	0.70	0.70	0.70	0.69	0.70	0.69	0.69	0.69
27	0.72	0.73	0.73	0.73	0.72	0.73	0.72	0.72	0.72
29	0.74	0.75	0.75	0.75	0.75	0.75	0.75	0.74	0.74
31	0.76	0.76	0.77	0.77	0.76	0.76	0.76	0.76	0.76
33	0.76	0.77	0.77	0.77	0.77	0.77	0.76	0.76	0.76
35	0.77	0.77	0.77	0.78	0.78	0.77	0.77	0.77	0.77
37	0.77	0.78	0.78	0.79	0.78	0.78	0.79	0.78	0.78
39	0.77	0.78	0.78	0.79	0.79	0.79	0.79	0.79	0.78
41	0.78	0.78	0.78	0.79	0.79	0.79	0.79	0.79	0.78

List of Publications

Lozano Vega, G., Benezeth, Y., Marzani, F., and Boochs, F. (2014b). Modular method of detection, localization, and counting of multiple-taxon pollen apertures using bag-of-words. *Journal of Electronic Imaging*, 23(5):053025.

Lozano Vega, G., Benezeth, Y., Marzani, F., and Boochs, F. (2014a). Analysis of relevant features for pollen classification. In Iliadis, L., Maglogiannis, I., and Papadopoulos, H., editors, *Artificial Intelligence Applications and Innovations*, volume 436 of *IFIP Advances in Information and Communication Technology*, pages 395–404. Springer Berlin Heidelberg.

Lozano Vega, G., Benezeth, Y., Marzani, F., and Boochs, F. (2013). Classification of pollen apertures using bag of words. In Petrosino, A., editor, *Image Analysis and Processing*, volume 8156 of *ILecture Notes in Computer Science*, pages 712–721. Springer Berlin Heidelberg.

Lozano Vega, G., Benezeth, Y., Uhler, M., Boochs, F., and Marzani, F. (2012). Sketch of an automatic image based pollen detection system. In *32. Wissenschaftlich-Technische Jahrestagung der DGPF*, volume Band 21, pages 202–209, Potsdam, Germany.

BIBLIOGRAPHY

Ahn, S. J., Rauh, W., and Warnecke, H.-J. (2001). Least-squares orthogonal distances fitting of circle, sphere, ellipse, hyperbola, and parabola. *Pattern Recognition*, 34(12):2283–2303.

Allen, G. (2006). An automated pollen recognition system. *Unpublished M. Eng. thesis, Massey University, Plamerston North, New Zealand*.

Allen, G., Hodgson, B., Marsland, S., Arnold, G., Flemmer, R., Flenley, J., and Fountain, D. (2006). Automatic recognition of light microscope pollen images. In *Image Vision and Computing New Zealand (IVCNZ 2006)*, pages 355–360. Massey University.

Bergmann, K. (2014). Aufkommen der allergologisch relevanten Pollen in Deutschland von 2001 bis 2013. Technical report, Stiftung Deutscher Polleninformationsdienst.

Borgefors, G. (1986). Distance transformations in digital images. *Computer vision, graphics, and image processing*, 34(3):344–371.

Boucher, A., Hidalgo, P. J., Thonnat, M., Belmonte, J., Galan, C., Bonton, P., and Tomczak, R. (2002). Development of a semi-automatic system for pollen recognition. *Aerobiologia*, 18(3-4):195–201.

Bousquet, J. et al. (2008). Allergic rhinitis and its impact on asthma (ARIA) 2008. *Allergy*, 63:8–160.

Buchner, R. and Weber, M. (2014). (2000 onwards). paldat - a palynological database: Descriptions, illustrations, identification, and information retrieval. www.paldat.org.

Byun, H. and Lee, S.-W. (2002). Applications of support vector machines for pattern recognition: A survey. In *Pattern recognition with support vector machines*, pages 213–236. Springer.

Cao, Y., Wang, C., Li, Z., Zhang, L., and Zhang, L. (2010). Spatial-bag-of-features. In *Computer Vision and Pattern Recognition, IEEE Conference on*, pages 3352–3359. IEEE.

Chabrier, S., Stoll, B., and Goujon, J.-B. (2012). SVM texture classification for tropical vegetation mapping. In *SPIE Asia-Pacific Remote Sensing*, pages 85270E–85270E. International Society for Optics and Photonics.

Chen, C., Hendriks, E. A., Duin, R. P., Reiber, J. H., Hiemstra, P. S., de Weger, L. A., and Stoel, B. C. (2006). Feasibility study on automated recognition of allergenic pollen: grass, birch and mugwort. *Aerobiologia*, 22(4):275–284.

Costa, C., Menesatti, P., Paglia, G., Pallottino, F., Aguzzi, J., Rimatori, V., Russo, G., Recupero, S., and Reforgiato Recupero, G. (2009). Quantitative evaluation of tarocco sweet orange fruit shape using optoelectronic elliptic Fourier based analysis. *Postharvest biology and Technology*, 54(1):38–47.

Crampton, J. S. (1995). Elliptic Fourier shape analysis of fossil bivalves: some practical considerations. *Lethaia*, 28(2):179–186.

Csurka, G., Dance, C., Fan, L., Willamowski, J., and Bray, C. (2004). Visual categorization with bags of keypoints. In *Workshop on statistical learning in computer vision, European Conference on Computer Vision*, volume 1, pages 1–22.

Culverhouse, P. F., Macleod, N., Williams, R., Benfield, M. C., Lopes, R. M., and Picheral, M. (2014). An empirical assessment of the consistency of taxonomic identifications. *Marine Biology Research*, 10(1):73–84.

Da Fontoura Da Costa, L. and Cesar Jr, R. M. (2000). *Shape analysis and classification: theory and practice*. CRC press.

D'Amato, G., Cecchi, L., Bonini, S., Nunes, C., Annesi-Maesano, I., Behrendt, H., Liccardi, G., Popov, T., and Van Cauwenberge, P. (2007). Allergenic pollen and pollen allergy in europe. *Allergy*, 62(9):976–990.

Dell'Anna, R., Lazzeri, P., Frisanco, M., Monti, F., Campeggi, F. M., Gottardini, E., and Bersani, M. (2009). Pollen discrimination and classification by Fourier transform infrared (FT-IR) microspectroscopy and machine learning. *Analytical and bioanalytical chemistry*, 394(5):1443–1452.

Duda, R. O. and Hart, P. E. (1972). Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15.

Dykewicz, M. S. and Hamilos, D. L. (2010). Rhinitis and sinusitis. *Journal of Allergy and Clinical Immunology*, 125(2):S103–S115.

Egerton, R. (2005). The scanning electron microscope. In *Physical Principles of Electron Microscopy*, pages 125–153. Springer US.

Erdtman, G. (1943). *An introduction to pollen analysis.* Waltham, Mass., the Chronica Botanica Co.; London, WIW Dawson and Sons, Ltd.

Fawcett, T. (2003). Roc graphs: Notes and practical considerations for researchers. Tech report, HPL-2003-4, HP Laboratories.

France, I., Duller, A., Lamb, H., and Duller, G. (1997). A comparative study of approaches to automatic pollen identification. In *Proceedings of the British Machine Vision Conference*.

Goto, S., Iwata, H., Shibano, S., Ohya, K., Suzuki, A., and Ogawa, H. (2005). Fruit shape variation in Fraxinus mandshurica var. japonica characterized using elliptic Fourier descriptors and the effect on flight duration. *Ecological Research*, 20(6):733–738.

Goyal, A. and Walia, E. (2014). Variants of dense descriptors and Zernike moments as features for accurate shape-based image retrieval. *Signal, Image and Video Processing*, 8(7):1273–1289.

Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, 46(1-3):389–422.

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, SMC-3(6):610–621.

Holt, K., Allen, G., Hodgson, R., Marsland, S., and Flenley, J. (2011). Progress towards an automated trainable pollen location and classifier system for use in the palynology laboratory. *Review of Palaeobotany and Palynology*, 167(3–4):175 – 183.

Huang, D., Shan, C., Ardabilian, M., Wang, Y., and Chen, L. (2011). Local binary patterns and its application to facial image analysis: a survey. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 41(6):765–781.

HuoRong, R., XinXin, Y., Yan, Z., Rui, C., JianWei, S., and Yang, L. (2013). Relative gradient local binary patterns method for face recognition under varying illuminations. *Journal of Electronic Imaging*, 22(4):043013–043013.

Institute of Plant Sciences, University of Bern (2003). The surface of microspores. www.botany.unibe.ch/paleo/pollen_e/surface.htm.

Iwata, H., Niikura, S., Matsuura, S., Takano, Y., and Ukai, Y. (1998). Evaluation of variation of root shape of japanese radish (Raphanus sativus L.) based on image analysis using elliptic Fourier descriptors. *Euphytica*, 102(2):143–149.

Jackson, S. T., Webb, R. S., Anderson, K. H., Overpeck, J. T., III, T. W., Williams, J. W., and Hansen, B. C. (2000). Vegetation and environment in eastern north america during the last glacial maximum. *Quaternary Science Reviews*, 19(6):489 – 508.

Joachims, T. (1998). Text categorization with support vector machines: Learning with many relevant features. In *European Conference on Machine Learning*, pages 137–142, Berlin. Springer.

Kawashima, S., Clot, B., Fujita, T., Takahashi, Y., and Nakamura, K. (2007). An algorithm and a device for counting airborne pollen automatically using laser optics. *Atmospheric Environment*, 41(36):7987–7993.

Kimme, C., Ballard, D., and Sklansky, J. (1975). Finding circles by an array of accumulators. *Communications of the ACM*, 18(2):120–122.

Koenderink, J. J. and van Doorn, A. J. (1987). Representation of local geometry in the visual system. *Biological cybernetics*, 55(6):367–375.

Kohavi, R. and John, G. H. (1997). Wrappers for feature subset selection. *Artificial Intelligence*, 97(1-2):273 – 324. Relevance.

Landsmeer, S. H., Hendriks, E. A., De Weger, L. A., Reiber, J. H., and Stoel, B. C. (2009). Detection of pollen grains in multifocal optical microscopy images of air samples. *Microscopy research and technique*, 72(6):424–430.

Lazebnik, S., Schmid, C., and Ponce, J. (2006). Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, volume 2, pages 2169–2178. IEEE.

Leopold, E. B., Birkebak, J., Reinink-Smith, L., Jayachandar, A. P., Narváez, P., and Zaborac-Reed, S. (2012). Pollen morphology of the three subgenera of alnus. *Palynology*, 36(1):131–151.

Li, P., Treloar, W., Flenley, J., and Empson, L. (2004). Towards automation of palynology 2: the use of texture measures and neural network analysis for automated identification of optical images of pollen grains. *Journal of quaternary science*, 19(8):755–762.

Lichtman, J. W. and Conchello, J.-A. (2005). Fluorescence microscopy. *Nature methods*, 2(12):910–919.

Lladó, X., Oliver, A., Freixenet, J., Martí, R., and Martí, J. (2009). A textural approach for mass false positive reduction in mammography. *Computerized Medical Imaging and Graphics*, 33(6):415–422.

López-Sastre, R. J., Tuytelaars, T., Acevedo-Rodríguez, F. J., and Maldonado-Bascón, S. (2011). Towards a more discriminative and semantic visual vocabulary. *Computer Vision and Image Understanding*, 115(3):415–425.

Mandrioli, P. (2000). Method for sampling and counting of airborne pollen and fungal spores. Technical report, Institute of Atmospheric and Oceanic Sciences (ISAO), National Research Council (CNR).

Mebatsion, H. K., Paliwal, J., and Jayas, D. S. (2012). A novel, invariant elliptic Fourier coefficient based classification of cereal grains. *Biosystems engineering*, 111(4):422–428.

Menesatti, P., Costa, C., Paglia, G., Pallottino, F., D'Andrea, S., Rimatori, V., and Aguzzi, J. (2008). Shape-based methodology for multivariate discrimination among italian hazelnut cultivars. *Biosystems Engineering*, 101(4):417–424.

Mhangara, P. and Odindi, J. (2013). Potential of texture-based classification in urban landscapes using multispectral aerial photos. *South African Journal of Science*, 109(3-4):1–8.

Mikolajczyk, K. and Schmid, C. (2005). A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(10):1615–1630.

Mildenhall, D., Wiltshire, P., and Bryant, V. (2006). Forensic palynology: Why do it and how it works. *Forensic Science International*, 163(3):163 – 172. Forensic Palynology.

Mitsumoto, K., Yabusaki, K., and Aoyagi, H. (2009). Classification of pollen species using autofluorescence image analysis. *Journal of bioscience and bioengineering*, 107(1):90–94.

Mullins, J. and Emberlin, J. (1997). Sampling pollens. *Journal of Aerosol Science*, 28(3):365–370.

Nanni, L., Lumini, A., and Brahnam, S. (2012). Survey on LBP based texture descriptors for image classification. *Expert Systems with Applications*, 39(3):3634–3641.

Negrini, A. (1992). Pollens as allergens. *Aerobiologia*, 8(1):9–15.

Neto, J. C., Meyer, G. E., Jones, D. D., and Samal, A. K. (2006). Plant species identification using elliptic Fourier leaf shape analysis. *Computers and electronics in agriculture*, 50(2):121–134.

Nixon, M. (2008). *Feature extraction & image processing*. Academic Press.

Nowak, E., Jurie, F., and Triggs, B. (2006). Sampling strategies for bag-of-features image classification. In *European Conference on Computer Vision*, pages 490–503. Springer.

O'Higgins, P. (1997). Methodological issues in the description of forms. *Fourier descriptors and their applications in biology*, pages 74–105.

Ojala, T., Pietikäinen, M., and Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59.

Ojala, T., Pietikäinen, M., and Mäenpää, T. (2000). Gray scale and rotation invariant texture classification with local binary patterns. In *European Conference on Computer Vision*, pages 404–420. Springer.

Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27.

Pelleg, D., Moore, A. W., et al. (2000). X-means: Extending k-means with efficient estimation of the number of clusters. In *International Conference on Machine Learning*, pages 727–734.

Perveen, A. (2006). A contribution to the pollen morphology of family Gramineae. *World Applied Sciences Journal*, 1(2):60–65.

Provost, F. and Fawcett, T. (2001). Robust classification for imprecise environments. *Machine Learning*, 42(3):203–231.

Punt, W., Hoen, P., Blackmore, S., Le Thomas, A., et al. (2007). Glossary of pollen and spore terminology. *Review of Palaeobotany and Palynology*, 143(1):1–81.

Ranzato, M., Taylor, P., House, J., Flagan, R., LeCun, Y., and Perona, P. (2007). Automatic recognition of biological particles in microscopic images. *Pattern Recognition Letters*, 28(1):31–39.

Réseau National de Surveillance Aerobiologique (2014). Les pollens: Principaux pollens allergisants. www.pollens.fr/le-reseau/les-pollens.php.

Rodriguez, Y. (2006). *Face detection and verification using local binary patterns*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne.

Rodríguez-Damián, M., Cernadas, E., Formella, A., and González, A. (2003). Automatic identification and classification of pollen of the Urticaceae family. In *Advanced Concepts for Intelligent Vision Systems, Proceedings of*, pages 38–45.

Rohlf, F. J. and Archie, J. W. (1984). A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Systematic Biology*, 33(3):302–317.

Ronneberger, O. (2007). *3D invariants for automated pollen recognition*. PhD thesis, Albert-Ludwigs-Universität Freiburg im Breisgau.

Ronneberger, O., Wang, Q., and Burkhardt, H. (2007). 3D invariants with high robustness to local deformations for automated pollen recognition. In *Pattern Recognition*, pages 425–435. Springer.

Saito, T. and Toriwaki, J.-I. (1994). New algorithms for euclidean distance transformation of an n-dimensional digitized picture with applications. *Pattern Recognition*, 27(11):1551 – 1565.

Scharring, S., Brandenburg, A., Breitfuss, G., Burkhardt, H., Dunkhorst, W., von Ehr, M., Fratz, M., Giel, D., Heimann, U., Koch, W., et al. (2006). Online monitoring of airborne allergenic particles (OMNIBUSS). *Biophotonics, edited by: Popp, J. and Strehle, M., WILEY-VCH, Weinheim*, pages 31–87.

Semwogerere, D. and Weeks, E. R. (2005). Confocal microscopy. *Encyclopedia of Biomaterials and Biomedical Engineering*, pages 1–10.

Sergios Theodoridis, K. K. (2006). Pattern recognition. In *Pattern Recognition*. Academic Press, San Diego, third edition edition.

Sivaguru, M., Mander, L., Fried, G., and Punyasena, S. W. (2012). Capturing the surface texture and shape of pollen: a comparison of microscopy techniques. *PloS one*, 7(6):e39129.

Soldevilla, C. G., González, P. C., Teno, P. A., and Vilches, E. D. (2007). *Spanish Aerobiology Network (REA): management and quality manual*. Servicio de Publicaciones, Universidad de Córdoba.

Suzuki, S. and Abe, K. (1985). Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing*, 30(1):32–46.

Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):411–423.

Tou, J. Y., Tay, Y. H., and Lau, P. Y. (2009). A comparative study for texture classification techniques on wood species recognition problem. In *Fifth International Conference on Natural Computation*, volume 5, pages 8–12. IEEE.

Truong, Q. H. (2013). *Knowledge-based 3D point clouds processing*. PhD thesis, Université de Bourgogne.

Unay, D. and Ekin, A. (2008). Intensity versus texture for medical image search and retrival. In *Biomedical Imaging: From Nano to Macro, 5th IEEE International Symposium on*, pages 241–244. IEEE.

Wang, B., Liu, Y., Xiao, W., Xu, W., and Zhang, M. (2013). Position and locality constrained soft coding for human action recognition. *Journal of Electronic Imaging*, 22(4):041118–041118.

Wayne, R. (2009). *Light and video microscopy*. Academic Press.

Williams, K., Munkvold, J., and Sorrells, M. (2013). Comparison of digital image analysis using elliptic Fourier descriptors and major dimensions to phenotype seed shape in hexaploid wheat (Triticum aestivum L.). *Euphytica*, 190(1):99–116.

Yoshioka, Y., Iwata, H., Ohsawa, R., and Ninomiya, S. (2004). Analysis of petal shape variation of Primula sieboldii by elliptic Fourier descriptors and principal component analysis. *Annals of Botany*, 94(5):657–664.

Zhang, E. and Mayo, M. (2010). Enhanced spatial pyramid matching using log-polar-based image subdivision and representation. In *Digital Image Computing: Techniques and Applications, International Conference on*, pages 208–213. IEEE.

Zhang, Y., Fountain, D., Hodgson, R., Flenley, J., and Gunetileke, S. (2004). Towards automation of palynology 3: pollen pattern recognition using gabor transforms and digital moments. *Journal of quaternary science*, 19(8):763–768.

Zulpe, N. and Pawar, V. (2012). GLCM textural features for brain tumor classification. *International Journal of Computer Science Issues*, 9(3).

List of Figures

List of Tables

Abstract:

The correct classification of airborne pollen is relevant for medical treatment of allergies, and the regular manual process is costly and time consuming. An automatic processing would increase considerably the potential of pollen counting. Modern computer vision techniques enable the detection of discriminant pollen characteristics. In this thesis, a set of relevant image-based features for the recognition of top allergenic pollen taxa is proposed and analyzed. The foundation of our proposal is the evaluation of groups of features that can properly describe pollen in terms of shape, texture, size and apertures. The features are extracted on typical brightfield microscope images that enable the easy reproducibility of the method. A process of feature selection is applied to each group for the determination of relevance.
Regarding apertures, a flexible method for detection, localization and counting of apertures of different pollen taxa with varying appearances is proposed. Aperture description is based on primitive images following the Bag-of-Words strategy. A confidence map is built from the classification confidence of sampled regions. From this map, aperture features are extracted, which include the count of apertures. The method is designed to be extended modularly to new aperture types employing the same algorithm to build individual classifiers.
The feature groups are tested individually and jointly on of the most allergenic pollen taxa in Germany. They demonstrated to overcome the intra-class variance and inter-class similarity in a SVM classification scheme. The global joint test led to accuracy of 98.2%, comparable to the state-of-the-art procedures

Keywords: pattern recognition, classification, feature extraction, feature selection, object extraction, bag of words palynology, apertures

Résumé :

Le traitement médical des allergies nécessite la caractérisation des pollens en suspension dans l'air. Toutefois, cette tâche requiert des temps d'analyse très longs lorsqu'elle est réalisée de manière manuelle. Une approche automatique améliorerait ainsi considérablement les applications potentielles du comptage de pollens. Les dernières techniques d'analyse d'images permettent la détection de caractéristiques discriminantes. C'est pourquoi nous proposons dans cette thèse un ensemble de caractéristiques pertinentes issues d'images pour la reconnaissance des principales classes de pollen allergènes. Le cœur de notre étude est l'évaluation de groupes de caractéristiques capables de décrire correctement les pollens en termes de forme, texture, taille et ouverture. Les caractéristiques sont extraites d'images acquises classiquement sous microscope, permettant la reproductibilité de la méthode. Une étape de sélection des caractéristiques est appliquée à chaque groupe pour évaluer sa pertinence.
Concernant les apertures présentes sur certains pollens, une méthode adaptative de détection, localisation et comptage pour différentes classes de pollens avec des apparences variées est proposée. La description des apertures se base sur une stratégie de type Sac-de-Mots appliquée à des primitives issues des images. Une carte de confiance est construite à partir de la confiance donnée à la classification des régions de l'image échantillonnée. De cette carte sont extraites des caractéristiques propres aux apertures, permettant leur comptage. La méthode est conçue pour être étendue de façon modulable à de nouveaux types d'apertures en utilisant le même algorithme mais avec un classifieur spécifique.
Les groupes de caractéristiques ont été testés individuellement et conjointement sur les classes de pollens les plus répandues en Allemagne. Nous avons montré leur efficacité lors d'une classification de type SVM, notamment en surpassant la variance intra-classe et la similarité inter-classe. Les résultats obtenus en utilisant conjointement tous les groupes de caractéristiques ont abouti à une précision de 98,2 %, comparable à l'état de l'art.

Mots-clés : reconnaissance de formes, classification, extraction de caractéristiques, sélection de caractéristiques, extraction d'objets, sac-de-mots, palynologie, apertures

SPIM

■ École doctorale SPIM - Université de Bourgogne/UFR ST BP 47870 F - 21078 Dijon cedex
■ tél. +33 (0)3 80 39 59 10 ■ ed-spim@univ-fcomte.fr ■ www.ed-spim.univ-fcomte.fr

UNIVERSITÉ DE BOURGOGNE