General-purpose Object Recognition in 3D Volume Data Sets using Gray-Scale Invariants – Classification of Airborne Pollen-Grains Recorded with a Confocal Laser Scanning Microscope

Olaf Ronneberger / Hans Burkhardt
Albert-Ludwigs-University of Freiburg
Computer Science Department
79110 Freiburg, Germany
Olaf@Ronneberger.net / burkhardt@informatik.uni-freiburg.de

Eckart Schultz
German Weather Service
Human Biometeorology
79104 Freiburg, Germany
Eckart.Schultz@dwd.de

Abstract

A technique is described which may be employed to establish a fully automated system for recognition of airborne pollen. As the different pollen taxa have only marginal differences, a full 3D volume data set of the pollen grain was recorded with a confocal laser scanning microscope (LSM) at a voxel size of about $(0.2\mu m)^3$. This represents an intrinsic and complete data set. 14 invariant gray-scale features based on an integration over the 3D Euclidian transformation group with nonlinear kernels were extracted from these volume data sets. The classification was done with support vector machines. The use of these general gray scale features allows to easily adapt the system to other objectives (e.g., pollen of a special area) or even other objects than pollen (e.g., spores, bacteria etc.) just by exchanging the reference data base. When using a reference data base with the 26 most important German pollen taxa (385 samples), the recognition rate is 92%. With a special database for allergological purposes recognizing only Corylus, Alnus, Betula, Poaceae, Secale, Artemisia *and "allergological non-relevant" the recognition rate is 97.4%.*

1 Introduction

About 10% of the human population are allergic to pollen. Today's pollen-forecasts are based on time consuming and expensive "manual" pollen countings done by experienced microscopists. Real-time data of actual pollen-concentration are not available by that technique. In contrast to the microscopist, a pollen-recognition system based on image recognition techniques could be integrated into a pollen-trap to provide such real-time data.

Even though pattern recognition on images is widely used in several biological applications, there are only very few papers in the literature dealing with pollen recognition and most of them focus on fossil pollen and 2D data [6, 7, 4]. Bonton et al.[1] proceed like human microscopists, i.e. they use multiple images of the pollen from different focus planes and extract the same "high-level" features as human do (e.g., the number of pores) by use of a plenty of highly pollen-taxa-specific or even pollen-grain-orientation-specific algorithms for feature extraction and they employ (to the authors knowledge) a simple hard-coded classifier. Using a set of 350 pollen grains from 30 different taxa the according recognition rate is about 73%.

In contrast to such approaches, we use 3D volume data and a general-purpose feature extraction, namely Euclidian gray-scale invariants [11, 3], and we employ a classifier (a set of Support Vector Machines [13]) that is trained automatically with a labeled reference data base. In our programs the only a priori assumption is, that the objects are rigid and have random orientations and positions. In other words the system can easily be adopted to objects other than pollen (e.g., spores, bacteria etc.) just by exchanging the reference data base.

2 Material and Methods

2.1 Sampling, Preparation and Recording

To set up a reference data base, the pollen grains were directly collected from the plants of interest in order to prepare pure samples of the following pollen taxa:

Acer, Artemisia, Alnus, Alnus viridis, Betula, Carpinus, Corylus, Chenopodium, Compositae, Cruciferae, Fagus, Quercus, Aesculus, Juglans, Fraxinus, Plantago, Platanus, Poaceae, Secale, Rumex, Populus, Salix, Taxus, Tilia, Ulmus, Urtica.

In conventional pollen counting translucent microscopy is used. The pollen recognition of pollen which have been
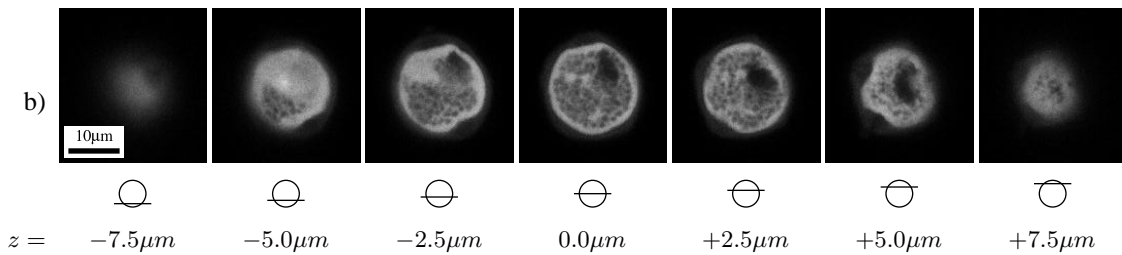
Figure 1. Alnus (alder) pollen grain: slices of the volume data set recorded with a confocal laser scanning microscope

collected in the open air, is complicated since one is confronted with a huge variety of particles of not only biological origin. The strong primary fluorescence of pollen provides an easily accessible feature which allows to reliably separate them from the background and the other mostly inorganic particles.

Even for a human pollen counter it is hard to recognize a pollen from a single 2D view at some unfavorable orientations of the pollen. As today's computer codes are still by far inferior to the object recognition capabilities of a human observer, the identification of all the pollen from a single 2D image is extremely unlikely[8].

To obtain sufficient information for the recognition the microscopist focuses the microscope to different planes of the pollen grain. Similarly we record 2D images from several focus planes and stack them up to a volume data set. Translucent microscopy is not well suited for this purpose, because the recorded images are the result of complicated integrals of light defraction and refraction due to the inhomogeneous refraction coefficient of the pollen grain and its surrounding. In fluorescence microscopy, however, all fluorescence active molecules of the pollen act as small light sources. The resulting image therefore can be regarded as the measurement of the local fluorescence activity, which is largely independent from the direction of viewing and the direction of illumination.

As a general problem, conventional imaging systems generate a superposition of the wanted image of the focused plane and out of focus images of neighboring planes. In order to eliminate the contribution from these non-focused planes, one can use either confocal microscopy, which eliminates unwanted light by hardware components providing images with the highest possible quality at very high costs, though. An alternative are deconvolution techniques (Wiener filter), which remove the light dispersion by post-processing of the digital images taken with a conventional fluorescence microscope. For the development of the recognition system we started with high-quality images from a confocal laser-scan microscope (Fig. 1)

2.2 Pattern Recognition with gray-scale invariants

A quite simple but very powerful way of a general feature extraction is the computation of so-called "gray-scale invariants", which were described first for two-dimensional image data [11, 3], but can be straightforwardly extended to three-dimensional volumetric data [10]. These gray-scale invariants do not need any segmentation within the object, but operate directly on the gray-values of the data set.

The advantageous property of such an invariant feature is the following: The set of all possible 3D volume data sets of one individual pollen grain – scanned in all possible positions and orientations (Euclidian motion) – is an equivalence class. An invariant transformation is able to map all elements of this equivalence class into one point of the feature space and there represents one bit of information on the intrinsic structure of the considered pollen grain, independent of its position and orientation (Fig. 2).
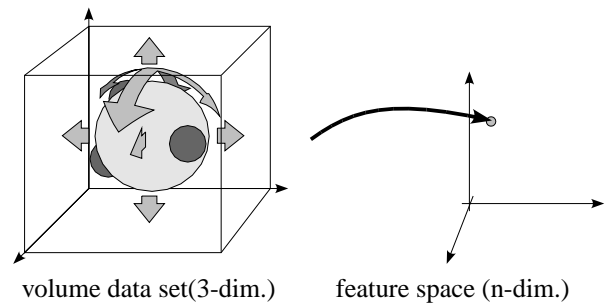

volume data set(3-dim.)   feature space (n-dim.)

Figure 2. Invariant transformation: All representations of an object (here: the object at any orientation or position) are mapped into the same point in the feature space.

The basic recipe for calculating these invariants is to take a non-linear kernel function of local support $f(\mathbf{X})$ in order to relate or combine the grey scale values of some neighboring pixels or voxels and to integrate the result of this function over all possible representations of the object in

the equivalence class [11].

$$T[f](\mathbf{X}) \quad := \quad \int_G f(g\mathbf{X})dg \qquad (1)$$

f : Kernel function
$\mathbf{X}$: gray-value image or volume data
G : transformation group
g : one element of the transformation group

For the sake of clearness, the 2D version of the formulas (for images) are presented in the following. By replacing the 2D translations and rotations with 3D operations and the images with volume data sets one obtains the 3D versions of the formulae:

For rigid objects, which is a fair approximation of pollen grains in the present context, the different elements of the equivalence class can be described by an Euclidian transformation (rotation and translation) of the object:

$$T[f](\mathbf{X}) := \int_{\vec{x}=0}^{\vec{x}_{max}} \int_{\varphi=0}^{2\pi} f(g_{\vec{x},\varphi}\mathbf{X})d\varphi d\vec{x} \qquad (2)$$

$\vec{x}_{max}$: extension of the image

Actually, it is not necessary to apply the transformation to the full image, instead the kernel function can be transformed which considerably speeds up the computation and results in linear complexity of the algorithm $O(N)$. This is illustrated by an example in figure 3.

A further speedup of this still expensive computation is accomplished for a special class of kernel functions by using a convolution with the image of a circle (or in 3D: of a spherical surface) $\mathbf{C}$. This convolution may be computed by means of the Fast Fourier Transform (FFT). For kernel functions of the type

$$f(\mathbf{X}) = f_a\Big(\mathbf{X}(\vec{0})\Big) \cdot f_b\Big(\mathbf{X}(\vec{q})\Big) \qquad (3)$$

f_a, f_b : any functions that transform the gray values
$\vec{q}$: span of the kernel function

one can rewrite equation 2 for the two-dimensional case using $\mathbf{A} := f_a(\mathbf{X})$ and $\mathbf{B} := f_b(\mathbf{X})$ as

$$T[f](\mathbf{X}) :=$$

$$\int\!\!\!\int_{\substack{x=0\\y=0}}^{\substack{N_x\\N_y}} \mathbf{A}(x,y) \int_{\varphi=0}^{2\pi} \mathbf{B}\left(x + |\vec{q}|\cos(\varphi), y + |\vec{q}|\sin(\varphi)\right) d\varphi \; dxdy$$
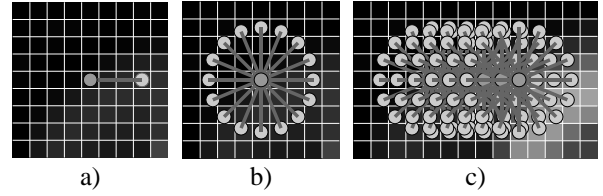

Figure 3. Calculation of a 2D gray-scale invariant: (a) non-linear kernel function for combining some neighboring pixels, e.g., the multiplication of two gray values of distance 3. (b) Evaluation for all angles. The results are summed up, to become invariant to rotations of the object. Gray values at fractional pixel-positions are bilinearly interpolated (c) Evaluation at all possible positions of the image. Again the results are summed up, to become invariant to translations of the object.

$$T = \sum \mathbf{X} \cdot (\mathbf{X} * \mathbf{C})$$
$$T = \sum \mathbf{X} \cdot \text{FFT}^{-1}\Big(\text{FFT}(\mathbf{X}) \cdot \text{FFT}(\mathbf{C})\Big)$$

Figure 4. Fast calculation of a special class of gray scale invariants: The sequential evaluation of the rotated kernel functions (as shown in fig. 3b) is split into two steps: step 1: the gray values touched by the second kernel point within the rotation are summed up. step 2: the result is multiplied with the gray value of the first kernel point. The evaluation of step 1 for all positions in the image is a simple convolution which could efficiently be calculated by means of the Fast Fourier Transform (FFT).

which then can be written also as

$$T[f](\mathbf{X}) \quad := \quad \int\!\!\!\int_{\substack{x=0\\y=0}}^{\substack{N_x\\N_y}} \mathbf{A}(x,y) \cdot (\mathbf{B} * \mathbf{C})\,(x,y)\; dxdy \qquad (4)$$

where $\mathbf{C}(x,y) = \begin{cases} 1 & : \quad \sqrt{x^2+y^2} = |\vec{q}| \\ 0 & : \quad \text{otherwise} \end{cases}$

and '$*$' denotes a convolution. This again is illustrated for one example in fig. 4. Besides saving computing time we are released from the decision, by which angular steps we

should proceed in the computation, which is not trivial, particularly in 3D volumes [10].

A more general method for saving computing costs has been described in [12]: the considered features are computed only approximately (Monte-Carlo-integration). Once the permissible error is fixed, this results in a constant complexity, independent from the size of the image.

Even though these features were designed to be only invariant to Euclidian transformations, they are also quite robust against other transformations like articulated motion or even slight topological deformations, due to the finiteness of the kernel support [3]

Different kernel functions can be used in order to develop a set of gray scale invariants which are adapted to a given problem. In fact, it is not difficult to construct features that provide a required discrimination power. Using a small-scale kernel results in a feature which is sensitive to small-scale structures of the object. For example coarse or fine-grained plasm. Correspondingly large-scale kernels sense the large-scale structure of the object, e.g., the difference between spherical and ellipsoid objects.

In the case of the pollen recognition, the following features turned out to provide a high discrimination performance: a vector of 14 features is constructed by evaluating the two kernel functions, $f(\mathbf{X}) = \mathbf{X}(0,0,0) \cdot \mathbf{X}(0,0,2)$ and $f(\mathbf{X}) = \sqrt{\mathbf{X}(0,0,0)} \cdot \sqrt{\mathbf{X}(0,0,2)}$ at 7 different scalings of the object (1:1, 1:2, 1:4, 1:8, 1:16, 1:32 and 1:64). Since the gray scale values of the input volume data sets were normalized to unit variance the elements of the feature vector are in the range $[-1 : 1]$ corresponding to normalized correlation coefficients.

Due to the non-linearity of the transformation and the particularly shape of the resulting clusters in the feature space, a simple MAP classifier based on normal distributions does not perform satisfactorily. A much better recognition rate is achieved by the so-called support vector machines [13]. The principal idea behind the support vector machine is to identify the clusters by searching for the thickest hyperplane, which separates this cluster from the remaining points. A good introduction to the theory of SVMs is given by C. J. Burges Tutorial [2]

2.3 Measuring the recognition rate

In order to measure the quality of our recognition system, we have used a reference data base with the 26 most relevant German pollen taxa. 3D volume data sets of about 15 samples from each pollen taxon were recorded with a resolution of ca. 5 voxels/μm in each direction using a confocal laser scanning microscope with a 40x oil-objective, an excitation wavelength of 450-490nm and an emission wavelength greater than 510nm.

With these 385 high-quality volume data sets we tested our recognition system using the "leave one out" technique. As classifier we use a set of 26 SVMs with a Gaussian kernel where each SVM was trained to separate one particular class from the rest. The radius of the Gaussian kernel was determined by optimizing the recognition rate.

3 Results and Discussion

The achieved recognition rate for all 26 taxa was about 92%. The details are listed in table 1. For pollen forecasts, however, we are interested only in the allergologically relevant pollen. So it doesn't matter if the computer cannot distinguish, for example, between an *Ulmus* and a *Platanus*. So we can put all the allergologically irrelevant taxa into one class resulting in a recognition rate for allergologically relevant pollen of 97.4%.
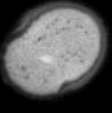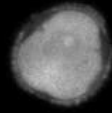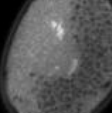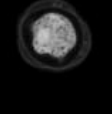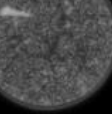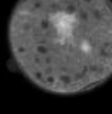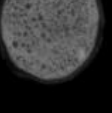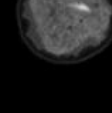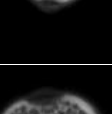
With regard to this high recognition rate one has to keep in mind that the examined pollen may have less variations in size and morphology than airborne pollen because the pollen for each taxa were taken just from one plant. Furthermore our reference pollen are not expected to have deformations due to sampling stress in the pollen trap [5] and there are no contaminated or agglomerated pollen grains.
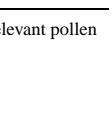
Anyway, the nearly perfect performance of the automatic recognition working on these high-quality pollen images is encouraging enough to test the technique with reduced data quality by using a normal fluorescence microscope with subsequent deconvolution and real-world air samples with deformed or contaminated pollen. Last not least, we can use a pollen-calendar to reduce the reference data base to the seasonally possible set of pollen, which again should increase the recognition rate.

For establishing this system in a laboratory environment, one main aspect is the time needed for the analysis. The imaging with the LSM currently takes about 40s per object (depending on its size) and the calculation of the 14 gray scale invariants for a 128^3 voxel volume takes about 15s on a Pentium II Dual-Processor PC with 400MHz, so that we end up with a recognition time of about 1min for each object. This time will be drastically reduced by using the conventional fluorescence microscope, which can record the same 3D volume in a few seconds. On the computational side, the use of a faster processor and a reduction of the resolution by a factor of 2 in each direction finally may reduce the recognition time to a few seconds per object.

Our current work also focuses on the 2D pre-recognition of the objects, so that only pollen with an unfavorable orientation or other doubtful objects have to be subjected to the relatively time-consuming 3D recognition.

Table 1. Classification Results using 3D LSM Data (leave-one-out Classification)

Acer
14 correct
1 → *Tilia*

Artemisia[a]
13 correct
1 → *Compositae*
1 → *Platanus*

Alnus[a]
15 correct
no wrong

Alnus viridis[a]
12 correct
no wrong

Betula[a]
13 correct
2 → *Plantago*

Carpinus
14 correct
no wrong

Corylus[a]
13 correct
1 → *Alnus*

Chenopodium
12 correct
1 → *Quercus*
1 → *Plantago*
1 → *Populus*

Compositae
15 correct
no wrong

Cruciferae
13 correct
1 → *Acer*
1 → *Populus*

Fagus
15 correct
no wrong

Quercus
11 correct
1 → *Acer*
2 → *Chenopodium*
1 → *Plantago*

Aesculus
15 correct
no wrong

Juglans
13 correct
1 → *Carpinus*
1 → *Poaceae*

Fraxinus
12 correct
2 → *Compositae*
1 → *Plantago*

Plantago
13 correct
2 → *Fraxinus*

Platanus
15 correct
no wrong

Poaceae[a]
15 correct
no wrong

Secale[a]
11 correct
3 → *Fagus*
1 → *Tilia*

Rumex
15 correct
no wrong

Populus
14 correct
1 → *Chenopodium*

Salix
15 correct
no wrong

Taxus
15 correct
no wrong

Tilia
14 correct
1 → *Poaceae*

Ulmus
12 correct
2 → *Platanus*
1 → *Populus*

Urtica
14 correct
1 → *Platanus*

[a] Allergological relevant pollen

4 Conclusions and Outlook

The utilization of fluorescence microscopy for the pollen recognition was found to have considerable advantages over translucent microscopy. It gives valuable information for the segmentation of the pollen from the background and it allows the use of 3D volume analysis straightforwardly, because one can measure the local properties of the object largely independently of its orientation towards the direction of viewing and illumination.

The 3D information is required for a highly reliable pollen recognition even by a microscopist, because a single 2D view of a pollen grain at an unfavorable orientation does not provide sufficient information for an unequivocal classification.

The selected gray scale invariants were found to be a powerful approach for object recognition on 3D volume data. Up to now, there is no pollen-specific code in our recognition software, which makes it reusable for a wide range of other problems. The only limitation is, that the objects have to be rigid within the support of the selected kernel functions.

It was demonstrated, that our system can recognize pollen on the basis of 3D volume data with a good reliability by using data recorded with a confocal laser scanning microscope and pollen, which are not deformed or contaminated. The recognition rate of 92% for all 26 pollen taxa and 97.4% when combining all allergological irrelevant in one class is encouraging. The result gives reason to test the procedure with reduced data quality by using a conventional fluorescence microscope with subsequent deconvolution and air samples with deformed or contaminated pollen. We expect, that the described approach may replace the tedious and time consuming work of manual pollen counting in near future.

Furthermore this approach opens up the possibility to directly integrate such a system into a pollen sampler resulting in an online pollen monitor.

More informations concerning our project including some 3D volume images of pollen may be found on our web pages [9].

Acknowledgements

References

[1] P. Bonton, A. Boucher, et al. Colour image in 2d and 3d microscopy for the automation of pollen rate measurement. In *Proceeding of 8th ECS and Image Analysis*, Bordeaux, France, 2001.

[2] C. J. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2:121–167, 1998.

[3] H. Burkhardt and S. Siggelkow. Invariant features in pattern recognition - fundamentals and applications. In C. Kotropoulos and I. Pitas, editors, *Nonlinear Model-Based Image/Video Processing and Analysis*, pages 269–307. John Wiley & Sons, 2001.

[4] I. France, A. Duller, H. Lamb, and G. Duller. A comparative study of model based and neural network based approaches to automatic pollen identification. In *British Machine Vision Conference*, pages 340–349, 1997.

[5] J. Hirst. An automatic volumetric spore trap. *Ann. Appl. Biol.*, 36:257–265, 1952.

[6] M. Langford, G. Taylor, and J. Flenley. The application of texture analysis for automated pollen identification. In *Proceedings of a Conference on Identification and Pattern Recognition*, pages 729–739, Toulouse, 1986. Univ. Paul Sabatini.

[7] M. Langford, G. Taylor, and J. Flenley. Computerized identification of pollen grains by texture analysis. *Review of Palaeobotany and Palynology*, 64:197–203, 1990.

[8] M. Mazière. étude de faisabilité pour la reconnaissance automatique de grains de pollen. http://www-sop.inria.fr/orion/Publications/, 1997. Rapport de DEA, Universit de Nice-Sophia Antipolis.

[9] O. Ronneberger. Homepage of the project "automatic identification and counting of airborne pollen grains". http://lmb.informatik.uni-freiburg.de/research, 2001.

[10] M. Schael and S. Siggelkow. Invariant grey-scale features for 3d sensor-data. In *Proceedings of the International Conference on Pattern Recognition (ICPR2000)*, pages 531 – 535, Barcelona, Spain, September 2000.

[11] H. Schulz-Mirbach. Invariant features for gray scale images. In G. Sagerer, S. Posch, and F. Kummert, editors, *17. DAGM-Symposium "Mustererkennung"*, Informatik aktuell, pages 1–14. Springer, 1995.

[12] S. Siggelkow and M. Schael. Fast estimation of invariant features. In W. Förstner, J. Buhmann, A. Faber, and P. Faber, editors, *Mustererkennung, DAGM 1999*, Bonn, sep 1999. Springer.

[13] V. N. Vapnik. *The nature of statistical learning theory*. Springer, 1995.