Nachrichtenbl. Deut. Pflanzenschutzd., **58** (11), S. 309–314, 2006, ISSN 0027-7479.

German Weather Service, Freiburg[1]; MeteoSwiss, Zurich[2]; Institute of Computer Sciences, University of Freiburg[3]; Fraunhofer Institute of Physical Measurement Techniques, Freiburg[4]; Fraunhofer Institute of Toxicology and Experimental Medicine, Hannover[5]; Helmut Hund GmbH, Wetzlar[6]; Breitfuss Messtechnik GmbH, Harpstedt[7]

Automatic pollen recognition – developments and perspectives

Automatische Pollenerkennung – Entwicklungen und Perspektiven

Stefan Scharring[1], Eckart Schultz[1], Ulrich Heimann[1], Regula Gehrig[2], Claudio Defila[2], Barbara Köhler[2], Hans Burkhardt[3], Olaf Ronneberger[3], Qing Wang[3], Albrecht Brandenburg[4], Gerd Sulz[4], Markus v. Ehr[4], Dominik Giel[4], Markus Fratz[4], Wolfgang Koch[5], Wilhelm Dunkhorst[5], Hubert Lödding[5], Werner Müller[6], Gernot Breitfuss[7]

Abstract

Automatic pollen recognition has been developed based on so-called gray-scale invariants, which characterise pollen grains independently from their position and orientation on the microscopic sample. Thus, pollen features can be extracted from the gray-scale images of transmitted light and fluorescence microscopy.

In a first step, this approach is demonstrated with *Ambrosia* pollen of samples from a Burkard sampler, where pollen are collected from ambient air on a sticky tape mounted on a slowly rotating drum. Self-learning Support Vector Machines create a classification model from the gray-scale invariants of the particles on three Burkard samples from Mezzana (Ticino), Switzerland. Automatic pattern recognition is tested with 13 other samples from the period between July, 20th and September, 9th 2004. A recall of 77.3 % has been found for the automatic recognition of *Ambrosia* pollen, together with a precision of 84.0 % for this classification. Falsely negative classified objects can partly be ascribed to agglomerated pollen, the number of falsely positive classified objects can be reduced by a more specific classification mode.

Automatic pollen recognition provides the basis for the development of a fully automated system that combines sampling, particle deposition onto a surface suitable for optical analysis, automatic preparation, microscopic imaging techniques, pattern recognition and the hourly output of number concentration of airborne pollen.

Key words: Automatic pollen recognition, *Ambrosia* pollen, primary fluorescence, gray-scale invariants, Support Vector Machines, pattern recognition, pollen monitor, Burkard sampler, online-measurement

Zusammenfassung

Die automatische Erkennung von Pollen wurde auf der Basis von so genannten Grauwert-Invarianten entwickelt, die Pollen unabhängig von ihrer Position und Orientierung auf der mikroskopischen Probe charakterisieren. Somit können Polleneigenschaften aus den mikroskopischen Graustufen-Bildern (Durchlicht und Fluoreszenz) abgeleitet werden.

Zunächst wird dieser Ansatz am Beispiel von *Ambrosia*-Pollen auf Proben aus Burkardfallen demonstriert, in denen Pollen aus der Außenluft auf einem Klebeband gesammelt werden, das auf einer langsam rotierenden Trommel befestigt ist. Selbstlernende Support-Vector-Machines erzeugen ein Klassifikationsmodell aus den Grauwert-Invarianten der Partikel von drei Burkardproben aus Mezzana (Tessin), Schweiz. Die automatische Mustererkennung wird an 13 weiteren Proben aus der Zeit zwischen dem 20. Juli und dem 9. September 2004 getestet. Für die automatische Erkennung von *Ambrosia*-Pollen ergibt sich eine Erkennungsrate von 77,3 % bei einer Bestimmungsgenauigkeit von 84,0 % für diese Klassifizierung. Falsch negativ klassifizierte Objekte können teilweise auf agglomerierte Pollen zurückgeführt werden, die Zahl falsch positiv klassifizierter Objekte kann durch eine spezifischere Klassifikation reduziert werden.

Die automatische Pollenerkennung bildet die Basis der Entwicklung eines voll automatisierten Systems, das Probenahme, Partikelabscheidung auf einer mikroskopierfähigen Oberfläche, automatische Präparation, mikroskopische Abbildungstechniken, Mustererkennung und die stündliche Angabe der Anzahlkonzentration luftgetragener Pollen vereint.

Stichwörter: Automatische Pollenerkennung, *Ambrosia*-Pollen, Eigenfluoreszenz, Grauwert-Invarianten, Support-Vector-Machines, Mustererkennung, Pollenmonitor, Burkardfalle, Online-Messung

1 Introduction

There is growing evidence that climate change might facilitate the geographical spread of particular plant species to new areas which become climatically suitable. But also effects of changes in land use, socio-cultural changes as well as international transport and tourism are obviously promoting the spread of plant species. The occurrence of some invasive species can result in particular risks for health and requires the control especially of such pollen characterised by high allergic potential e.g. Cupressaceae, ragweed or mugwort, even at very low pollen concentrations.

For conventional analysis, pollen are sampled from ambient air with instruments like the Burkard sampler that provide samples for subsequent microscopic analysis of pollen deposited on a sticky carrier. Pollen counting under the microscope is done by eye. This work is a demanding and time consuming task even for experienced microscopists. From the samples, daily average pollen concentration is derived in routine measuring networks.

This method of pollen analysis includes some disadvantages concerning sufferers from allergy:

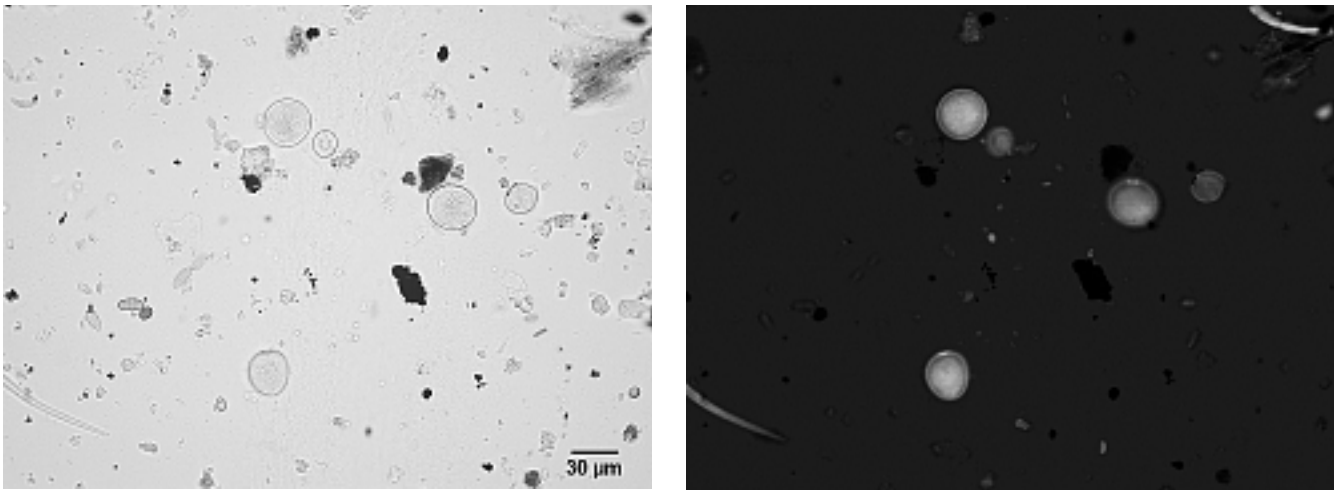

Fig. 1. Pollen grains are rather translucent and difficult to detect on an aerosol sample without specific staining (left). In the case of automatic pollen recognition the primary fluorescence light of pollen is used for their detection (right).

- There is a strong dependency of the pollen concentration on the current meteorological situation (MAKRA, 2004). However, the daily pollen concentration data are not available earlier than the day after sampling. Thus, the precision of the pollen forecast is strongly limited.
- Because the capacity of visual pollen analysis is usually limited to a sampling area representing 1 m³ air volume/day, the statistic error belonging to the determination of low pollen concentrations (which are still relevant for species like *Ambrosia*) is considerably high.
- The pollen concentration and thereby the allergic stress can change significantly during the day. This temporal course is usually not determined within routine analysis due to the high effort in visual microscopy.
- Furthermore, the quality and reliability of routine data may vary according to qualification and commitment of the pollen counters.

Thus, there is a strong public demand to improve the pollen forecast by using pollen concentration data that exhibit the following features:

- real-time information,
- higher temporal resolution and
- high quality, specified by reproducibility, measurement uncertainty, detection limit, recall and precision.

Automatic pollen recognition is the first step towards an online pollen information service and an objective pollen forecast. This recognition method will be explained in chapter 2 and demonstrated with *Ambrosia* pollen in section 3.1. In the following section 3.2, the integration of pollen sampling, sample preparation, automated microscopic analysis and pollen recognition into an online-monitor will be described that can work as a stand-alone-system for field measurements providing hourly online pollen data (SCHARRING, 2006).

2 Methods

2.1. Fluorescence microscopy

Pollen grains are rather translucent and unobtrusive among the variety of aerosol particles, e.g., mineral dust, tire abrasion, plant fragments and fungal spores that can be found in an aerosol sample. For the visual determination and counting of pollen a special staining like safranin or fuchsin is used to highlight the pollen from the background and to facilitate pollen counting.

Automatic pollen analysis is capable to set the staining procedure aside by using the primary fluorescence of pollen: If pollen grains are excited with ultraviolet light they emit green fluorescence light, cf. Fig. 1 (TAYLOR, 1989). This specific property can be used to separate the pollen from the background (segmentation).

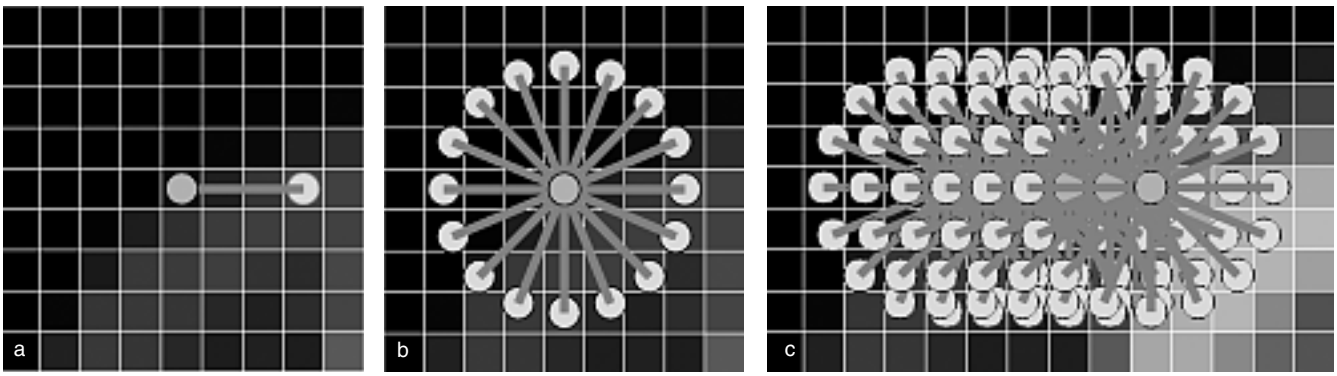

Fig. 2. Calculation of a 2D gray-scale invariant: (a) Selection of a non-linear kernel function for combining some neighboring pixels, in this example the combination of two gray values of distance 3. (b) This kernel function is evaluated for all angles and the results are summed up, to become invariant to rotations of the object. (c) This set of rotated kernel functions is evaluated at all possible positions of the image and the results are summed up, to become invariant to translations of the object. The result obtained is independent from the angle and position of the object in the image.

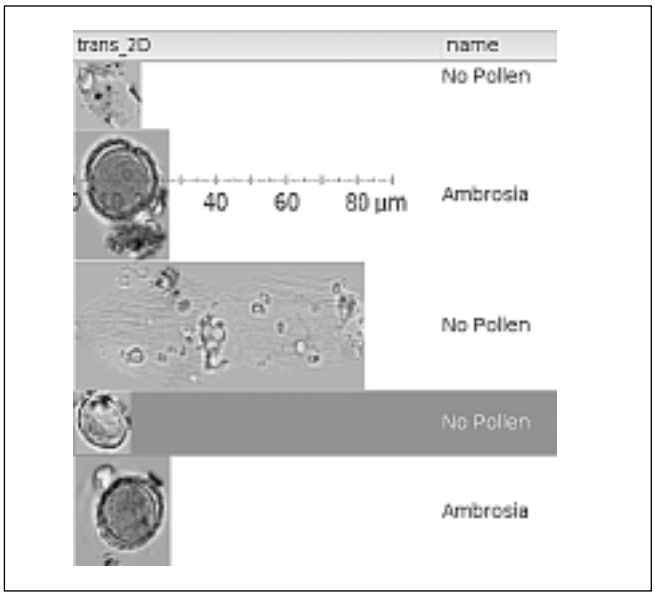

Fig. 3. Screenshot of the pollen file manager software developed for the labelling of pollen.

2.2. Pattern recognition

While conventional methods of pattern recognition focus on morphologic features and try to imitate the human recognition process, the method developed for pollen recognition employs a special feature of digitised images, the lightness of each pixel, described by its gray-value (0 to 255). This method is explained in Figure 2 (BURKHARDT, 2001; RONNEBERGER, 2002).

The kernel function that links the gray-values of each two pixels determines the resulting value of the invariant. As an example, the multiplication of the gray-values of each two pixels according to Figure 2 will usually lead to a different result than their addition. Hence, depending on the choice of the kernel function, different invariants can be derived from the image data that characterise the unknown object. A set of n invariants of an object can be regarded as a feature vector in the n-dimensional feature space. Feature vectors of similar objects form a cluster in the feature space, e.g. an *Ambrosia* cluster and a cluster of other aerosol particles, because the resulting values of the corresponding invariants are similar. Support Vector Machines can determine the interfaces that separate these clusters from each other and can be used as a self-learning tool for the classification of unknown objects (VAPNIK, 1995).

3 Results

3.1. Laboratory analysis

In this section, the automatic pollen analysis of Burkard samples is described. The sampler was located in Mezzana, Switzerland, in the vicinity of the Italian border. A period with high concentrations of *Ambrosia* pollen has been investigated (August, 27th to September, 9th 2004). Additionally, two samples from days with low concentrations of *Ambrosia* pollen have been selected (July, 20th and August, 20th of 2004).

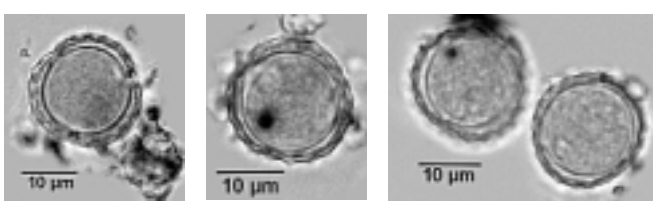

Fig. 4. Microscopic images (20f. magnification) of *Ambrosia* pollen with interfering dust (left, middle) and agglomerated *Ambrosia* pollen (right).

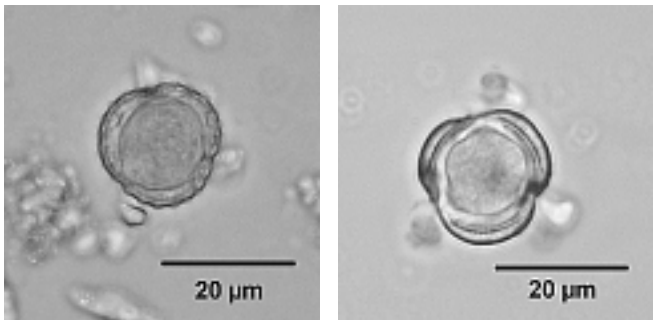

Fig. 5. Microscopic images (40f. magnification) of *Ambrosia* pollen (left) and *Artemisia* pollen (right).

3 of these samples have been chosen to built up a reference for the automatic recognition of the remaining 13 samples. On the three reference samples all objects exhibiting fluorescence properties were segmented from the microscopic images. Overall 6347 objects have been found, among them 408 containing *Ambrosia* pollen and 387 with different pollen, mostly *Urtica* pollen. Many different objects with primary fluorescence have been found like plant abrasion and a lot of particles showing poor fluorescence light. For classification purposes all objects have simply been divided in two classes: *Ambrosia* pollen and other objects, labelled by experienced microscopists, cf. Fig. 3.

Labelling the objects and the calculating their invariants enables the Support Vector Machines to learn to distinguish between *Ambrosia* pollen and other objects. A classification model has been derived from the clusters of the feature vectors belonging to the objects within the reference. This model was applied to classify the objects on the remaining 13 samples. The automatic recognition was compared with human classification, the results are shown in Table 1.

While pattern recognition of pollen is an ambitious task, dealing with real world samples is the crucial challenge for a successful pollen monitoring. Figure 4 depicts the main problems that occur on real world samples.

Table 1. Recognition results for *Ambrosia* pollen on Burkard samples. The percentage of falsely negative classified *Ambrosia* pollen is given by (1 – recall) whereas the percentage of objects classified falsely positive as *Ambrosia* pollen is given by (1 – precision)

Particle class	Number	Recall	Precision
Ambrosia pollen	538	77.3 %	84.0 %
Other particles	44 083	99.8 %	99.7 %

Table 2. Occurrence and classification of other pollen on the Burkard samples

Pollen species	Number	Classified as *Ambrosia*
Urtica	325	0
Graminae/Poaceae	71	0
Chenopodium	67	1
Castanea	53	0
Humulus	37	1
Artemisia	26	13
Plantago	21	0
Rumex	13	1
Compositae	12	7
Others	1084	8

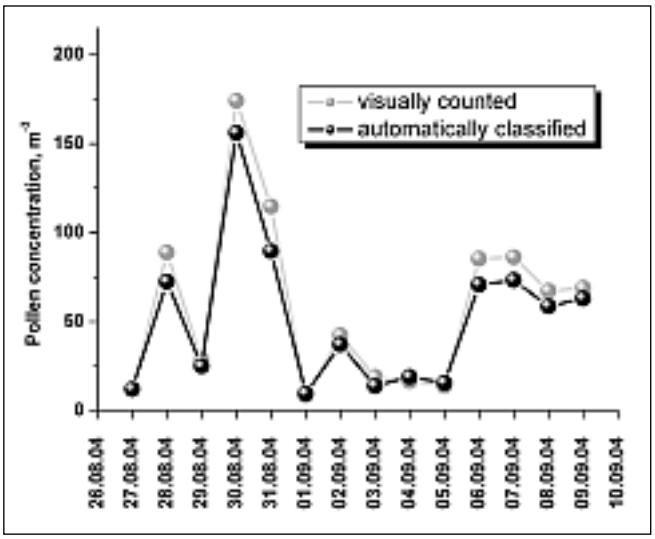

Fig. 6. Temporal course of *Ambrosia* pollen concentration in Mezzana (Ticino), Switzerland, results from visual counting vs. automatic pollen recognition.

Pollen grains are mostly located close to other atmospheric dust particles on the sample. These particles comprise a great morphologic variety and may disturb the automatic pollen recognition.

Furthermore, 102 of 538 *Ambrosia* pollen were agglomerated in pairs (82), triplets (12) or quadruplets (8) respectively. Only 3 of these 47 agglomerates have been recognised to be an object containing *Ambrosia* pollen – without a specification of the number of pollen within the agglomerate. Therefore, it was calculated how the recognition rate would increase if the agglomerated pollen were neglected. In this case, the recall would increase from 77.3 % up to 84.3 %. Thus it is obvious that the recognition of *Ambrosia* can be optimised by improvements in segmentation of the agglomerates by the means of image analysis.

Furthermore, a more specific classification of other pollen would lead to a more precise recognition, cf. Tab. 2 and Fig. 5. If all other pollen were identified correctly and did not interfere with the recognition of *Ambrosia* pollen, the precision would increase from 84.0 % up to 89.3 %.

Nevertheless, the combination of falsely negative classified *Ambrosia* pollen and objects falsely positive classified as *Ambrosia* may lead to a classification result that is closer to the reality in the end. In this case, the overall number of objects classified as *Ambrosia* pollen amounts 91.8 % of the total number of *Ambrosia* pollen.

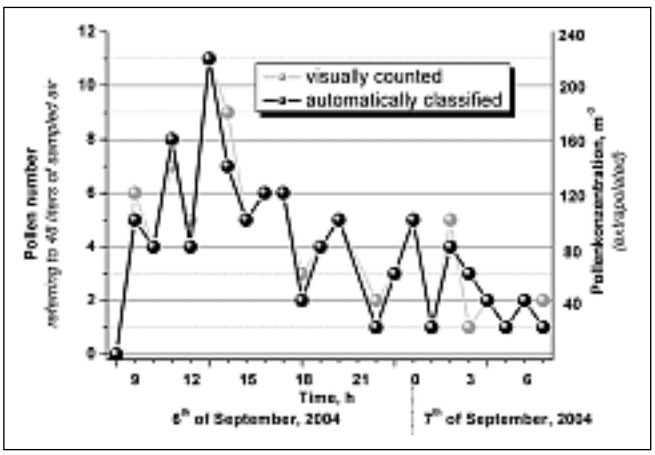

Fig. 7. Daily course of *Ambrosia* pollen concentration in Mezzana (Ticino), Switzerland, results from visual counting vs. automatic pollen recognition.

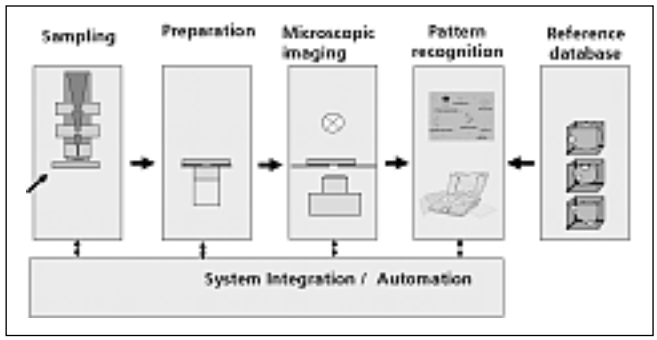

Fig. 8. Scheme of the steps from sampling to recognition within the pollen monitor.

Figure 6 shows the classification results of the two period (including samples from the reference).

Because the drum inside the Burkard sampler slowly rotates, the location of the particle on the sample allows a rough estimation about its sampling time. Automated microscopic analysis provides an easy access to these data and delivers results with higher temporal resolution (Fig. 7).

3.2. Online-monitor

The main disadvantage of laboratory analysis of pollen is the late availability of pollen data due to the time consuming process of sampling, transport, manual preparation, sample analysis. In-situ measurements providing online-analysis could give the chance to allow a more accurate pollen forecast considering the actual meteorological situation.

In the year 2003 the OMNIBUSS project (online monitoring of airborne allergenic particles) was raised. A demonstrator for the online-monitor comprising automated sampling, preparation, microscopic analysis and pattern recognition was presented in March 2005, see Figure 8.

In the sampling unit, a relatively large stream of outside air (Q_{in} = 70 m³/h) is sucked in through an inlet system designed for representative sampling also of the large particles up to an

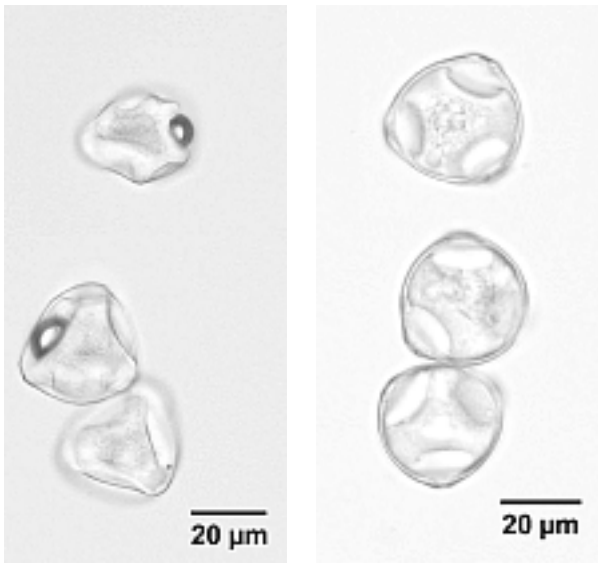

Fig. 9. Microscopic image (40f. magnification) of dried-out hazel pollen (left) and hazel pollen after rehydration in a polar liquid (right) respectively.

aerodynamic diameter of 60 µm. Only a side stream of Q_s = 6 m³/h, taken isokinetically from the overall sampling flow, is further processed for particle identification and counting. This is the relevant flow rate for the determination of the pollen concentration. The flow Q_s, however, can still be much larger than the maximum flow rate Q_d of the final step of particle deposition. This requires an additional conditioning unit (virtual impaction) in which the flow rate is sufficiently reduced without decreasing at the same time the flux of the relevant particles to be analysed. In the instrument Q_d = 2 l/min was chosen for the flow rate in the deposition unit. Hence, the impaction velocity is quite low (1.7 m/s) compared with Burkard samplers (6.0 m/s) providing a reduced mechanical stress for the impacted particles.

The lower cut-off of the impactor is set to around 10 µm aerodynamic diameter. This is a significantly higher cut-off compared with Burkard samplers (around 2.5 µm) and provides a great advantage. Disturbances of the pattern recognition with fine dust particles, as described above, can be avoided. Nevertheless, the lower cut-off can be set to 2.5 µm by the means of electrostatic precipitation, e.g. if the analysis of airborne fungal spores is required.

The sampling is followed by a short preparation process. As a sample carrier, a metal plate with circular holes is used, covered on its backside with a transparent cover slide for microscopic analysis. The resulting shallow cavities are filled to a certain level with a mixture of glycerine, gelatine, water and a surfactant. This mixture serves as an adhesive for particle sampling. Instead of adding an additional embedding liquid after sampling, the sampled particles are immersed by the same medium that serves for deposition in the sampling unit. The immersion is achieved by heating the sample to about 90 °C. At this temperature the medium is of lower viscosity so that the particles become immersed. While airborne, pollen dry out and change their shape and size due to the occurring water loss. Pollen recognition requires the re-establishment of their original shape. This is achieved by the up-take of a polar liquid like water or glycerine that the pollen grains acquire from the embedding liquid, cf. Fig. 9. This process typically takes only a few seconds.

After the automatic preparation process particles are located in quite different levels of the embedding medium. To obtain a full 3D volumetric data set of all particles image stacks for each contrasting technique steps of 2.5 µm distance are taken at each position of the sample area.

High power light emitting diodes (LED) are used for both transmitted light as well as fluorescence imaging. They have several advantages concerning lifetime, power consumption, heat generation, power stability and spectral stability in comparison to conventional used bulbs and high pressure arc lamps.

The system comprises sample carriers for a one week automatic operation stored in a stacker. By means of a fully computer controlled three axis motorised stage, carriers are transferred between the stacker and the sampling unit respectively the preparation unit and the specifically adapted microscopic unit. Although using only one motorised stage the system allows a simultaneous sampling of particles onto one carrier and a microscopic analysis at another, by which a nearly continuous collection and microscopic imaging of aerosol particles is possible with only short interruptions during transfer of carriers. Transfer time of the carriers is in the range of half a minute between the stacker and the other units of the system. The sampling period is intended to be one hour, resulting in one hour for subsequent sample preparation, analysis and notification of recognition results.

Table 3. Recognition results for *Alnus, Corylus,* and *Taxus* pollen on samples from the online-monitor

Particle class	Number	Recall	Precision
Alnus pollen	122	92.6 %	85.0 %
Corylus pollen	199	88.4 %	98.9 %
Taxus pollen	338	86.4 %	98.6 %
Other particles	283	97.5 %	82.4 %
Total	942	91.0 %	91.0 %

Fig. 10. Sampling site at the German weather service in Freiburg with the pollen monitor (left) and the Burkard sampler (right) for the field test 2006.

Preliminary tests have been carried out to assess the automatic pollen recognition with automatically prepared samples from the online-monitor, cf. Tab. 3. In this case, pollen grains have been added afterwards to aerosol samples of a period exhibiting very low pollen concentrations. As shown in Table 3, good recognitions rates can be achieved.

4 Discussion

A method for the automatic recognition of pollen was tested with Burkard samples. The quality of the automatic recognition of *Ambrosia* pollen can be characterised by recall (77.3 %), precision (84.0 %) and overall deviation from the true value (–8.2 %). These results are promising and show the capability of the self-learning pattern recognition.

The number of falsely negative classified objects can be reduced by using more specific algorithms for the segmentation process especially of agglomerated pollen. These algorithms and further tests with different invariants are subject to actual developments. On the other side, the number of falsely positive classified pollen can be reduced by taken more different pollen species into account for the reference database.

Automatic sampling, preparation, microscopy and recognition have been integrated in a system with the capacity to stand-alone-operation and online-data transfer. The samples of the pollen monitor from the sampling period 2005 are now used to create a reference comprising at least 20 different pollen species. This reference will be applied to the field test of the monitor in Freiburg i.Br. taking place in 2006 (Fig. 10).

Following the field test, for the years from 2007–2010, a limited-lot production of the monitor and the set-up of a pollen measurement network at meteorological stations of the German Weather Service is planned.

5 Acknowledgements

The contribution of the project partners and the financial support by the BMBF in the Biophotonik framework under grant No. 13N8368, 13N8367 13N8487, 13N8437 and 13N8372 are gratefully acknowledged.

6 Literature

BURKHARDT, H., S. SIGGELKOW: Invariant features in pattern recognition – fundamentals and applications. In: KOTROPOULOS, C., I. PITAS, ed.: Nonlinear Model-Based Image/Video Processing and Analysis, pages 269–307, Chichester, John Wiley & Sons, 2001.

MAKRA, L., M. JUHÀSZ, E. BORSOS, R. BÉCZI, 2004: Meteorological variables connected with airborne ragweed pollen in southern hungary. Int. J. Biometeorol. **49**, 37–47.

RONNEBERGER, O., E. SCHULTZ, H. BURKHARDT, 2002: Automated Pollen Recognition using 3D Volume Images from Fluorescence Microscopy. Aerobiologia **18**, 107–115.

SCHARRING, S. et al.: Online monitoring of airborne allergenic particles (OMNIBUSS). In: POPP, J., M. A. STREHLE, ed.: Biophotonics: Vision for better health care, Weinheim, Wiley-VCH, 2006, 31–86.

TAYLOR, D. L., Y.-L. WANG, ed., 1989: Fluorescence Microscopy of Living Cells in Culture. Methods in Cell Biology, **29** and **30**, San Diego Academic Press Inc.

VAPNIK, V. N.: The nature of statistical learning theory, 2nd ed., Berlin, Springer, 2000, 314 p.

Corresponding author: Uwe Kaminski, Deutscher Wetterdienst, Stefan-Meier-Str. 4-6, 79104 Freiburg, Germany, phone: +49(0)76 12 82 02 52, E-Mail: uwe.kaminski@dwd.de