Image Acquisition and Analysis of Hazardous Biological Material in Air

Christoph Sklarczyk[1], Horst Perner[2], Hans Rieder[1], Walter Arnold[1], and Petra Perner[2]

[1] Fraunhofer-Institute for Non-Destructive Testing (IZFP), Bldg. E 3.1 University, D-66123 Saarbrücken, Germany
[2] Institute of Computer Vision and Applied Computer Sciences, IBaI, Leipzig, Germany

Abstract. Human beings are exposed every day to bio-aerosols in the various fields of their personal and/or professional daily life. The European Commission has rules protecting employees in the workplace from biological hazards. Airborne fungi can be detected and identified by an image-acquisition and interpretation system. In this paper we present recent results on the development of an automated image acquisition, probe handling and image-interpretation system for airborne fungi identification. We explain the application domain and describe the development issues. The development strategy and the architecture of the system are described and some results are presented.

Keywords: Microscopic image acquisition, microbiological probe handling, image analysis, image interpretation, case-based object recognition, case-based reasoning.

1 Introduction

Airborne microorganisms are ubiquitously present in various indoor and outdoor environments. The potential implication of fungal contaminants in bio-aerosols on occupational health is recognized as a problem in several working environments. There is a concern on the exposure of workers to bio-aerosols especially in composting facilities, in agriculture, and in municipal waste treatment. The European Commission has therefore guiding rules protecting employees in the workplace from airborne biological hazards. In fact, there are an increasing number of incidents of building-related sickness, especially in offices and residential buildings. Some of these problems are attributed to biological agents, especially in relation to airborne fungal spores. However, the knowledge of health effects of indoor fungal contaminants is still limited. One of the reasons for this limitation is that appropriate methods for rapid and long-time monitoring of airborne microorganisms are not available.

Besides the detection of parameters relevant to occupational and public health, in many controlled environments the number of airborne microorganisms has to be kept below the permissible or recommended values, e.g. in clean rooms, in operating

P. Perner and O. Salvetti (Eds.): MDA 2006/2007, LNAI 4826, pp. 1–14, 2007.

theaters, and in domains of the food and pharmaceutical industry. Consequently, the continuous monitoring of airborne biological agents is a necessity for the detection of risks of human health as well as for the flawless operation of technological processes.

At present a variety of methods are used for the detection of fungal spores. The culture-based methods depend on the growth of spores on an agar plate and on the counting of colony-forming units [14]. Culture-independent methods are based on the enumeration of spores under a microscope, the use of a polymerase chain reaction or on DNA hybridization for the detection of fungi [14]. However, all these methods are limited by time-consuming procedures of sample preparation in the laboratory. This paper describes the development and the realization of an automated image-acquisition and probe handling unit of biologically dangerous substances and the automated analysis and interpretation of microscope images of these substances.

In the system described here, contaminated air containing bio-aerosols is collected in a defined volume via a carrier agent. They are recorded by an image-acquisition unit, counted, and classified. Their nature is determined by means of an automated image-analysis and interpretation system. Air samples are automatically acquired, prepared and transferred by a multi-axis servo-system to an image-acquisition unit based on a standard optical microscope with a digital color camera. This part of the system is described in Section 2. To obtain a sufficient image quality, special requirements have to be fulfilled by the image-acquisition unit which will be described in Section 3.

The variability of the biological objects is very broad. Given the constraints of the image acquisition, this variability is found in the appearance of the objects as well. There are no general features allowing one to discern the type of the detected fungi. In the system employed here, images are stored, and a more generalized description for the different appearances of the same objects is used. We will describe this novel case-based reasoning approach for the image analysis and its interpretation in Section 4. Finally, we summarize our work in Section 5.

2 System Requirements

The system to be developed should allow to collect dust and biological aerosols in well-defined volumes over microscope slides, deposit them there, image them with an appropriate method and count and classify them with an automated image analysis and interpretation method, in order to determine the following parameters from the images:

- Total number of airborne particles
- Classification of all particles according to the acquired image features
- Classification of biological particles, e.g. spores, fragments of fungal mycelia, and fragments of insects
- Number of respirable particles
- Total number of airborne particles of biological origin
- Number of dead particles of biological origin
- Number of viable and augmentable particles of biological origin

- Identification of species or geni exploiting the characteristic shapes of spores and pollen
- Proportion of airborne abiotic and biotic particles
- Proportion of dead and viable airborne microorganisms.

At the beginning of the project the following requirements concerning the optical and the mechanical system were defined:

- Color images should be produced in order to facilitate the separation of dead and living objects.
- It should be possible to generate images in at least three defined depths of field.
- A marker liquid like lactophenol should be used to further enhance the separation of dead and living objects (blue color for living objects). For that a cover slip is necessary in order to uniformly distribute the marker drop on the object slide.
- The object slide should be covered with an adhesive in order to fix the airborne germs.

Table 1. Strains of fungi used and selected properties of spores

Species	Strain no.	Spore shape	Spore color	Spore size [µm]
Alternaria alternata	J 37 (A[1])	Septated, clavate to ellipsoidal	Pale brown	18 – 83 × 7-18
Aspergillus niger	i400 (B[2])	Spherical, ornamented with warts and spines	Brown	Ø 3.5 - 5
Rhizopus stolonifer	J 07 (A)	Irregular in shape, often ovoid to elliptical, striate	Pale brown	7-15 × 6-8
Scopulariopsis brevicaulis	J26 (A)	Spherical to ovoid	Rose-brown	5-8 × 5-7
Ulocladium botrytis	i171(B)	Septated, ellipsoidal	Olive-brown	18-38 × 11-20
Wallemia sebi	J 35 (A)	Cubic to globose	Pale-brown	Ø 2.5 – 3.5

[1](A): from culture collection of JenaBios GmbH, Jena, Germany.
[2](B): from the fungal stock collection of the Institute of Microbiology, University of Jena, Jena, Germany.

Six fungal strains representing species with different spore types were identified as important species in different environments (Tab. 1) by our industrial project partner JenaBios GmbH. A database of images from the spores of these species was produced and was the basis of our development. The number of imaged spore per species was about 30-50. Since no commercial system was known fulfilling all requirements, a corresponding system was developed which is described in what follows.

3 The Automated Imaging System

3.1 The Microscopic Image-Acquisition System

Following the specifications given in Section 2 we developed an automated probe-handling and digital image-acquisition system for taking microbiological material from air samples [12]. An existing optical Leitz microscope was upgraded and expanded in its hardware. A lens from Olympus with a magnification of 60X and a numerical aperture of 0.7 was used. Its focal length of 1.7 mm provided sufficient clearance between the lens and the object slide including the cover glass to avoid collisions due to their variability in thickness. The lens was inserted in an autofocusing device from Physik Instrumente (PI, Karlsruhe, Germany) which was mounted on the lens revolver. A motorized xy-table from Märzhäuser (Wetzlar, Germany) with a motion controller was used to arbitrarily shift the object slide in both x and y direction. For the digital image acquisition a 1.4 Mpixel color digital camera from Soft Imaging System (SIS, Münster, Germany) was used. Our estimates showed that a pixel number larger than 1.4 Mpixel is sufficient for the given magnification. Fig. 1 demonstrates that the optical resolution is sufficient to recognize details in spores like Ulocladium.

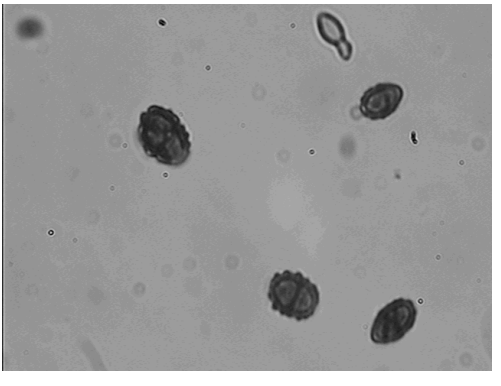

Fig. 1. Image demonstrating the resolution of the optical microscope used. The microscopical image displays spores of Ulocladium. The field of view is 134×100 µm². The sample was prepared by AUA/JenaBios, lens Olympus 60X/0.70. The resolution in this image is 5 µm.

The functions of image acquisition and image storage, movement of the specimen in x and y direction, and auto-focusing in z-direction are controlled by the AnalySIS Pro software from SIS. A pattern of images at any image position can be freely programmed and stored in a macro-code. This holds as well for the number of images to be captured. If necessary it is possible to capture automatically images at different depths of focus around the optimum position. By the automatic shading correction, the effect of an inhomogeneous illumination of the object can be removed.

3.2 The Automatic Probe-Acquisition and Handling System

The following chapter describes the main units and functions of the demonstration set-up realized in the course of the project. A stock of special object slides covered with a sticky layer from Umweltanalytik Holbach [1], (Fig. 2) is kept in a slide storage. A sliding gripper takes the lowest slide in the storage and transports it into the slit impactor from Umweltanalytik Holbach (Fig. 3). The object slides are separated by distance holders with a corresponding recess, in order to avoid sticking between the slides. The distance holder is removed by the same gripper, now moving in opposite direction and depositing the distance holder into a box. The distance holders can be used again when the slide deposit is reloaded.

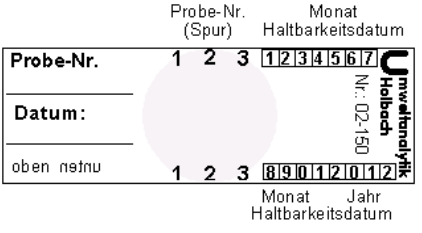

Fig. 2. Object slide of standard size 76×26×1 mm³ with a central sticky layer [1]; Image obtained from Umweltanalytik Holbach

Fig. 3. Slit impactor for collection of airborne particles [1]; Image from Umweltanalytik Holbach

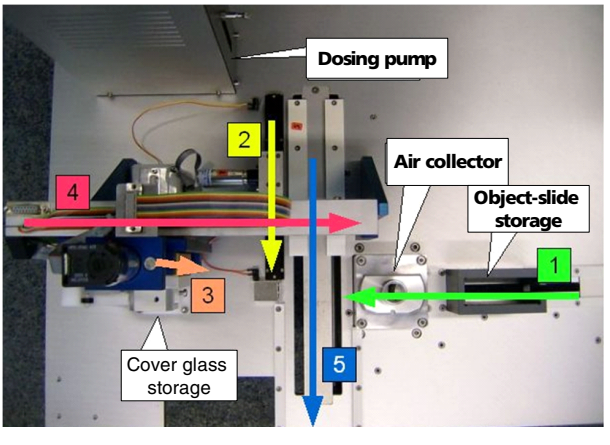

Fig. 4. Top view of the mechanical unit for moving object slides, indicating also the position of the cover-glass storage, the dosing pump for lactophenol, the slit impactor or air collector, and the storage for the object slides. The numbers 1 – 5 indicate the sequences of the movements; axis No. 6 is not shown.

In the slit impactor the air (Fig. 3), potentially containing airborne germs, is guided on the sticky area of the object slide by the air stream generated by a microprocessor controlled air pump. After a few tens of seconds which can be adjusted accordingly, the pump is switched off and the object slide is transported to the pipetting unit driven by the dosing pump (Cavro XL 3000 from Tecan Systems San Jose, Ca, USA. To this aim it has to change its transporting axis and thus its direction of movement. From a thin nozzle one drop of lactophenol is deposited on the sticky area of the object slide which is afterwards transported via the axis crossing to the cover-slip gripper unit. This gripper acts as a low-pressure sucker and takes one cover glass from the deposit and puts it with one edge first on the object slide. Then the cover glass falls down on the object slide and flattens the drop so that it will be distributed all over the sticky area forming a thin layer. In this way the airborne germs collected in the sticky layer are immersed in the lactophenol. In lactophenol living germs get a blue color. The object slide is then transported back to an axis crossing-point where it again changes its direction of movement by 90° and is transported to the xy-table of the microscope which takes over the slide and transports it directly under the lens. The timing of the transportation units, the air and dosing pump is controlled by a distributed multi-axis motion-unit. To this end an additional module was integrated into the AnalySIS Pro software. It controls the manual or automated shift of the xy-table between the image-acquisition position under the lens and the loading position, where the object slide is shifted from the object-slide preparation unit to the xy-table. After the object slide has reached the image acquisition position, the microscope camera then grabs the images at the programmed slide positions after auto-focusing of the microscope lens at each position. The cycle of shifting the xy-table to the defined positions, auto focusing, image acquisition and storage is programmable in a macro-code integrated into the AnalySis Pro software. This can also be done for other procedures like shading correction or image acquisition at different z-positions. After having finished the imaging sequence, the slide is transported away from the xy-table with a special arm

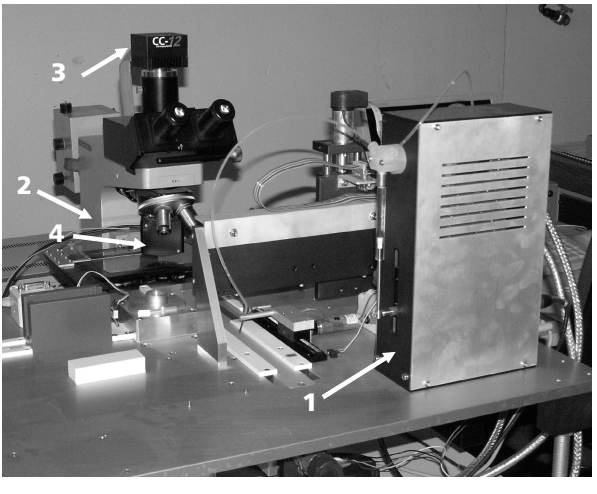

Fig. 5. Prototype set-up showing the dosing pump (arrow 1), several axes, the optical microscope with xy-table (arrow 2), and the digital camera (CC-12, arrow 3). The auto-focusing unit holds the lens (arrow 4).

and falls into a box. When the image grabbing procedure by the microscope unit is still under way, the object-slide preparation unit already starts with the preparation of a new object slide.

The object-slide preparation and manipulation is performed by a hardware controller and by dedicated software written in C++. The transfer from the AnalySIS Pro software to the C++ software and vice versa is controlled by a communication protocol as interface between both software units. Altogether six different mechanical axes have to be handled, not counting the axes of the xy-table (Fig. 4). The unit for object-slide preparation and the expanded microscope are shown in Fig. 5.

4 Image Analysis

Once an image has been taken it is given to the image-analysis unit for further processing. We describe the overall architecture of the system [4][5] and its single components in the next sections.

4.1 The Architecture

The architecture of the system is shown in Figure 6. Objects are recognized in the microscopic image by a case-based object-recognition unit [3]. This unit has a case-base of shapes (case base_1) for fungi spores and determines on a similarity-based inference if there are objects in the image that have a similar shape as the ones stored in the case base. In this case the objects get labeled and are transferred for further processing to the feature-extraction unit. To ensure proper performance of this unit, the general appearance of the shapes of the fungi spores have to be learned. To this end we have developed a semi-automated procedure [3] that allows one to acquire the shape information from the raw image data and to learn groups of shape-cases and general shape-cases. A more detailed description of the case-based object-matching unit can be found in Section 4.2.

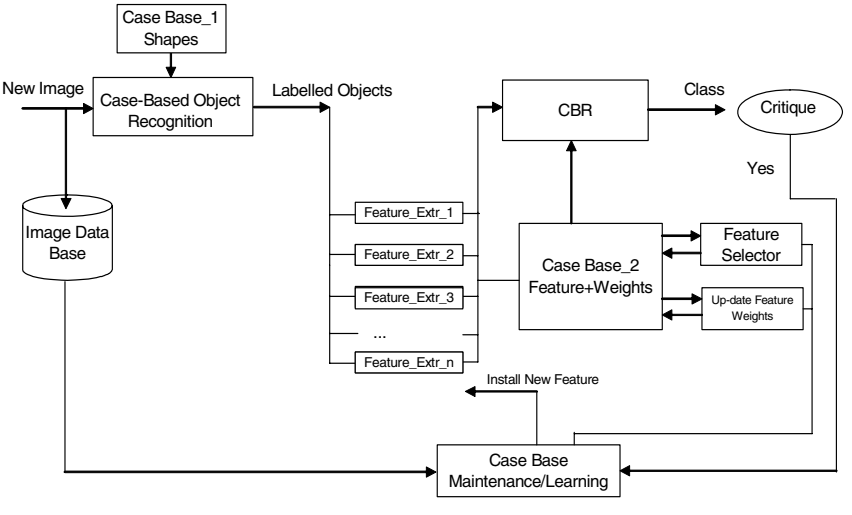

Fig. 6. System architecture

The feature-extraction procedures are based on the knowledge of an expert. Note that a particular application requires special feature descriptors. Therefore not all possible feature-extraction procedures can be implemented into such a system from the beginning. Our aim was to develop a special vocabulary and the associated feature-extraction procedures for application on fungi identification, as described in Section 4.3.

Based on the feature description, the second case-based reasoning unit decides about the type of the fungi spore. This unit employs a prototype-based classifier [11]. It starts its performance on prototypical cases that were selected or created by the expert. It can learn with time the different appearances of the fungi spores. The special features of this unit ensure its proper performance. It can learn the relevant prototypes from the subjectively selected set of prototypes, as well as create new prototypes. It can also learn the importance of the features of the cases. The final result of the system will be the identification of the fungi spores that appear in the image and the number of these spores. This is shown on the display of the system and in a file, together with the date and the time when the data were acquired.

Suppose that fungi species are wrongly identified by the system. Then a case-based maintenance process will start. First it has to be checked by the system developer whether new features have to be acquired for each case, or whether the whole case representation should be updated based on the learning procedures. The feature weights are learnt, as well as a subset of relevant features (see Section 4.4). To acquire new features means that the necessary feature-extraction procedures have to be developed and that for all cases the new features have to be calculated and fed into the existing case description. Therefore we keep the digital images acquired so far in the image-data base. Then the case representation has to be updated as well as the index structure. This ensures that we can come up step-by-step with a system which can describe the variability of the different biological objects that can appear.

4.2 Case-Based Object Recognition

The objects in the image are highly structured. Our study has shown that the images specified in Table 1 cannot be segmented by thresholding. The objects in the image may be occluded touching, or overlapping. It can also happen that only some parts of the objects appear in the image. Therefore we decided to use a case-based object recognition procedure [3] for the detection of objects in the image.

A case-based object-recognition method uses cases that generalize the original objects and matches them against the objects in the image. During this procedure a score is calculated that describes the quality of the fit between the object and the case. The case can be an object model which describes the inner appearance of the object as well as its contour. In our case the appearance of the entire objects can be very diverse. The shape seems to be the feature that generalizes the objects. Therefore, we decided to use contour models. We do not use the gray values of the model, but instead use the object's edges. For the score of the match between the contour of the object and the case we use a similarity measure based on the scalar product. It measures the average angle between the vectors of the template and the object.

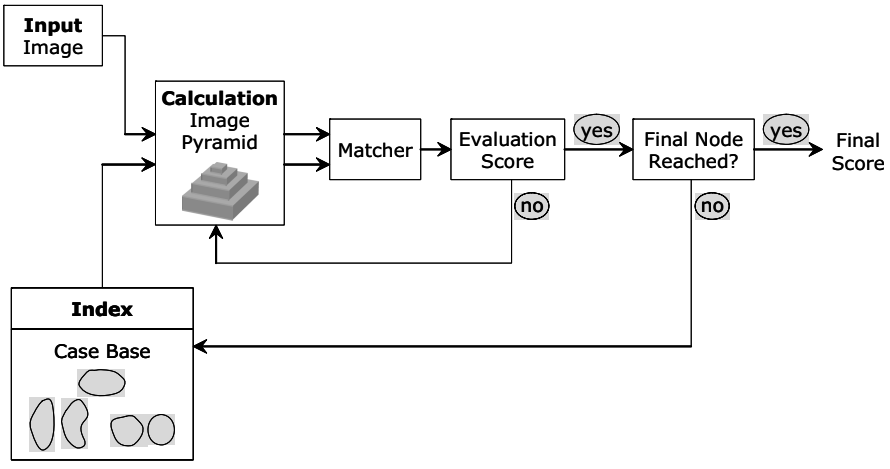

Fig. 7. Principle of case-based object-recognition architecture

4.2.1 Case-Base Generation

The acquisition of the case is done semi-automatically. Prototypical images are shown to an expert. The expert manually traces the contour of the object with the help of the cursor of the computer. Afterwards the number of contour points is reduced for data-reduction purposes by interpolating the marked contour by a first-order polynom. The marked object shapes are then aligned by the Procrustes Algorithm [4]. From the sample points the direction vector is calculated. From a set of shapes general groups of shapes are learnt by conceptual clustering which is a hierarchical incremental clustering method [5]. The prototype of each cluster is calculated by estimating the mean shape [5] of the set of shapes in the cluster and is taken as a case model.

4.2.2 Results for Case-Based Object Recognition

We had a total of 10 images for each class at our disposal. From this set of images two images were taken for the case generation. In these two images there were approx. 60 objects. These objects were labeled and taken for the case generation according to the procedure as described in Section 4.2.1. The result was a data base of cases. These cases were applied to the image for the particular class.

The threshold for the score was set to 0.8. We calculated the recognition rate as the number of objects that was recognized in the image to the total number of objects in the images. Note that the recognition rate can be higher than 100 %, since our procedure also operates in image regions where no objects are present due to background noise. The aim is to set-up the case-based object-recognition unit in such a way that the number of false alarms is low.

The results of the matching process are shown in Figs. 8 and 9. The highest recognition rate can be achieved for the objects Aspergillus niger and Scopularioupsi, since the shape of these objects does vary much. This is also expressed by the number of models, see Table 2. These classes have the lowest number of cases. For those classes where the variation of the shape of the objects is high, the number of the

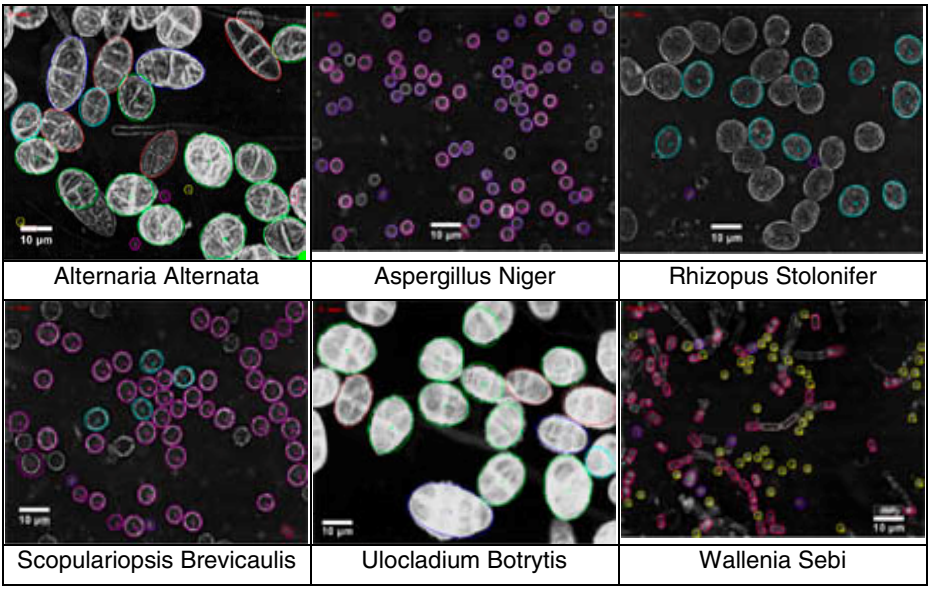

Fig. 8. Recognized objects in the image

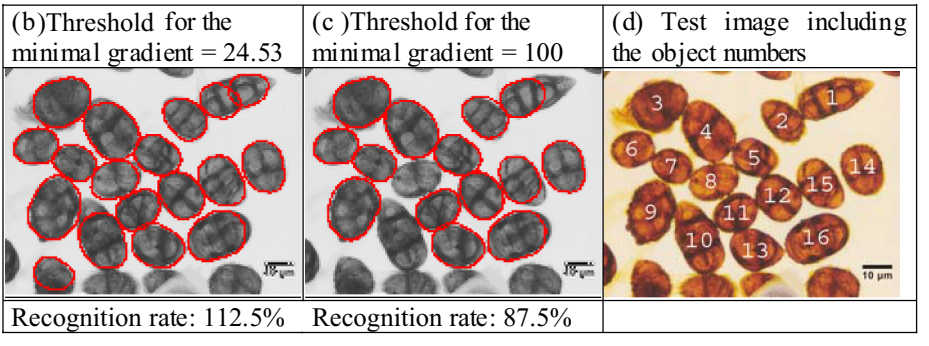

Fig. 9. Comparison of the matched objects by applying different thresholds for the minimal gradient

cases is also high. The recognition rate shows that we do not have enough cases to recognize the classes with a good recognition rate (see Ulocladium botrytis and Alternaria alternata). Therefore we need to increase the number of cases. For this task we developed an incremental procedure for the case acquisition in our tool. Objects that have not been recognized well will be displayed automatically for tracing and then the similarity to all other shapes will be calculated. The clustering will be done in an incremental fashion as well [5]. This procedure will ensure that we can learn the natural variation of the shape during the usage of the system.

Table 2. Results of matching

Classes	Number of models	Recognition rate
Alternaria alternata	34	65.9
Aspergillus niger	5	95.2
Rhizopus stolonifer	22	87.7
Scopularioupsi	8	94.5
Ulocladium botrytis	30	77.2
Wallenia sebi	10	90.3

4.3 Case Description and Feature Extraction

We choose an attribute-value pair-representation for the case description. The case consists of the solution which is the type of fungi spores and the features describing the visual properties of the object (see Figure 9). From each recognized object a set of features is extracted. One feature is the case number which represents the shape of the object, the similarity score between the actual shape and the shape in the case base, the size of the object, various gray-level features, and the texture inside the object. For the description of the texture we use our texture descriptor based on random sets described in [6].

4.4 Classification

Our case-based reasoning procedure to recognize spores relies on prototypical-based classification schemes [11]. Usually such schemes are generalized from a set of single cases. Here, we have prototypical cases represented as images that were selected by humans. That means when building our system, we start from the top and have to collect more information about the specific class during the usage of the system. Since a human has selected the prototypical images, his decision on the importance of an image might be biased, and to select only one image might be difficult for a human. He can have stored more than one image as prototypical images. Therefore we need to check the redundancy of the many prototypes for one class before taking them all into the case base. According to this consideration, our system has the following function to fulfill:

- Classification based on the nearest neighbor rule
- Prototype selection by a redundancy-reduction algorithm; Feature weighting to determine the importance of the features for the prototypes
- Feature-subset to select the relevant features from the whole set of the respective domain.

The classification method is based on the nearest-neighbor rule. Since the prototypes are available at the same time, we choose a decremental redundancy-reduction algorithm proposed by Chang [**7**] that deletes prototypes as long as the classification accuracy does not decrease. The feature-subset selection is based on the wrapper

approach [8] and an empirical feature-weighting learning method [9] is used. Furthermore, cross validation is used to estimate the classification accuracy. The prototype selection, the feature selection, and the feature-weighting steps are performed during each run of the cross-validation process. This rule classifies x in the category of its nearest neighbor [10]. More precisely, we call $x'_n \in \{x_1, x_2, ..., x_i, ..., x_n\}$ a nearest neighbor to x if $\min d(x_i, x) = d(x'_n, x)$, where i = 1, 2, …n. The nearest neighbor rule chooses to classify x into category C_n where x'_n is the nearest neighbor to x and x'_n belongs to class C_n. For the k-nearest neighbor we require k-samples of the same class to satisfy the decision rule. As a distance measure we use the Euclidean distance. The recognition rate was evaluated on a data base of 50 samples for each class based on cross-validation. The result is shown in Table 3. From that we can conclude that the classification accuracy is higher than the recognition rate for some classes. That means that it is more difficult to recognize the objects that are most likely to be fungi spores than to classify them based on the extracted features.

Table 3. Classification accuracy

Classes	Classification accuracy
Alternaria Alternata	90.4
Aspergillus Niger	95.0
Rhizopus stolonifer	92.0
Scopularioupsi	96.0
Ulocladium botrytis	94.0
Wallenia sebi	92.0

Fig. 10. Screenshot of the final system

A print-out of a result obtained by the system described in this paper is shown in Fig. 10. In the display the operator will find the acquired image in one window and in the other window the determined fungi spores and their total number. The system called Fungi PAD correctly identified the name of the fungi spores and their number.

5 Conclusion

In this paper a system for an automated image acquisition and analysis of hazardous biological material in air is described. It consists of an image-acquisition unit, its sample-handling hardware, and the image-interpretation system. The sample-handling and image-acquisition unit collects the airborne germs, deposits them on an object slide, disperses them with a marker fluid, and takes digital images of the germs in a programmable pattern. The stored images are analyzed in order to identify the germs based on a novel case-based object-recognition method. The case generation is done semi-automatically by manually tracing the contour of the object, by automated shape alignment and by shape clustering, and eventually by prototype calculation. Based on the acquired shape cases, the object-recognition unit identifies objects in the image that are likely to be fungi spores. The further examination of labeled objects is done by calculating more distinct object features, from which a prototype-based classifier determines the kind of fungi spores. After all objects have been classified by their type, the number of one type of fungi spores is calculated and displayed for the operator on the computer screen.

The recognition rate is good enough for on-line monitoring of environments. The final information can be used to determine its contamination with biological hazardous material. It can be used for health monitoring as well as for process control. The described system is the base.

References

1. Mücke, W., Lemmen, Ch.: Schimmelpilze-Vorkommen, Gesundheitsgefahren, Schutzmaßnahmen, Ecomed, Landsberg (2004)
2. Sklarczyk, C., Rieder, H., Rabe, U., Reiß, P., Santraine, B., Lonsdorfer, P.: Entwicklung von Methoden und Verfahren für die bildgebende Erfassung und rechnergestützte Auswertung biologischer Gefahrstoffe, IZFP-Bericht 060618-TW zum Projekt, Optische Bildnahmevorrichtung"; BMWi-Förderkennzeichen 16IN0148
3. http://www.umweltanalytik-holbach.de/index.html
4. Perner, P., Günther, T., Perner, H., Fiss, G., Ernst, R.: Health Monitoring by an Image Interpretation System - A System for Airborne Fungi Identification. In: Perner, P., Brause, R., Holzhütter, H.-G. (eds.) ISMDA 2003. LNCS, vol. 2868, pp. 64–77. Springer, Heidelberg (2003)
5. Perner, P.: Entwicklung von neuartigen Bildanalyse und -auswerteverfahren für luftgetragene biologische Gefahrstoffe, Schlussbericht IBaI, Teilvorhaben, BIOVISION, BMWI- Förderkennzeichen 16IN0147
6. Perner, P., Perner, H., Jänichen, S.: Recognition of Airborne Fungi Spores in Digital Microscopic Images. J. Artificial Intelligence in Medicine AIM 36, 137–157 (2006)
7. Perner, P.: Prototype-Based Classification. Int. J. of Applied Intelligences (in print, 2007)
8. Dryden, I.L., Mardia, K.V.: Statistical Shape Analysis. John Wiley & Sons, Chichester (1998)
9. Jaenichen, S., Perner, P.: Conceptual Clustering and Case Generalization of two-dimensional Forms. Computational Intelligence 22, 178–193 (2006)
10. Perner, P., Perner, H., Müller, B.: Mining Knowledge for Hep-2 Cell Image Classification. Journal Artificial Intelligence in Medicine 26, 161–173 (2002)

11. Chang, C.-L.: Finding Prototypes for Nearest Neighbor Classifiers. IEEE Trans. on Computers C-23, 1179–1184 (1974)
12. Perner, P.: Data Mining on Multimedia Data. LNCS, vol. 2558. Springer, Heidelberg (2002)
13. Wettschereck, D., Aha, D.W.: ”Weighting Features”, in Case-Based Reasoning Research and Development. In: Aamodt, A., Veloso, M.M. (eds.) Case-Based Reasoning Research and Development. LNCS, vol. 1010, pp. 347–358. Springer, Heidelberg (1995)
14. Aha, D.W., Kibler, D., Albert, M.A.: Instance-based Learning Algorithm. Machine Learning 6, 37–66 (1991)