CASE-BASED REASONING (CBR) IN PUBLIC HEALTH DELIVERY: AIRBORNE FUNGI IDENTIFICATION

Ibharalu, F.T.[1] and **Dawodu, G.A.**[2]
[1]Department of Computer Sciences, University of Agriculture, Abeokuta
[2]Department of Statistics, University of Agriculture, Abeokuta

Abstract

Human beings are exposed to aerial micro-organisms in the various fields of their personal and, or professional daily life. Some advanced communities have 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 a work on building an automatic cbr-based image interpretation system for the identification of different kinds of airborne fungi. We explain the application domain and describe the architecture of the cbr-based image interpretation system and conclude by explaining some of the advantages that a public health worker can derive from the proposed system in the cause of providing the masses with dependable health-care services.

Keywords: aerial microorganism, case-based reasoning (CBR), contaminant and public health.

Introduction

Airborne microorganisms are sometimes present around the various fields of indoor and outdoor environments. The potential implication of fungal contaminants in various aspects of occupational health is recognized as a problem in several working environments. There is concern on exposure of workers to infectious airborne microorganisms especially in composting facilities, in agriculture and in municipal waste treatment. At the moment, it is assumed that, in some industrial estates in Nigeria, 50% of their workplaces have negative influence by biological agents. Take Lagos, for instance, the streets of some local governmental areas contain overflowing rubbish bins, which usually pollutes the environment and during this process the release of airborne fungi cannot be ruled-out. Pollution of this nature will potentially endanger about 8 million employees and residents of the affected areas adversely. This phenomenon is better tackled squarely because if a scourge or epidemic can potentially ravage approximately 8 million of the estimated 12 million inhabitants of a state then the entire state is as good as being endangered. The European Commission has helped to set the appropriate standards by releasing guiding rules protecting employees in the workplace from airborne biological hazards. Of course, with respect to dirty environments there should be more awareness campaigns to warn people of the underlying hazards in their dirty habits.

Despite the release of the said guiding rules, there are an increasing number of incidents for building-related sickness, especially in office and residential building. Some of these problems are attributed, to the presence of biological agents, especially to airborne fungal spores. However, the knowledge of health effects of indoor fungal contaminants is still restricted since 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 and this can only be achieved through adequate, monitoring of proper waste management, and awareness campaign (e.g. clean rooms, operating theaters, domains of the food and pharmaceutical industry) (Aamodt *et al* 1998).

The continuous monitoring of airborne biological agents is consequently a necessity, so that the detection of risks to human health can be achieved early and the smooth sequence of technological processes can be sustained.

At present a variety of methods for detection of fungal spores is frequently used. The culture-based methods depend on growth of spores on an agar plate and counting of colony forming units. Culture-independent methods are based on the enumeration of spores under a microscope or use the polymerase chain reaction or DNA hybridization for the detection of fungi. However, all methods are limited by time-consuming procedures of sample preparation in the laboratory.

The aim of this paper is to develop and realize an image-acquisition unit of biologically dangerous substances and the automatic analysis of these images. The basic principle of the system shall lie in the fact, that dust and "bio-aerosols" are conducted in defined volumina via special carrying agents and separated there, that they are registered by an image-acquisition unit, counted, classified and that their nature is defined, by means of an automatic image-interpretation system.

The variability of the biological objects is very broad and the constraints of the image acquisition cause a broad variability in the appearance of the objects. Generalization about the objects can not be done at hand rather each case that appears in practice should be stored into the system and the system should learn more generalized description for the different appearances of the same objects over time. All that has been explained suggests to an experienced person the taking of a case-based reasoning approach for the image interpretation rather than a generalized approach.

In this paper we present our initial work on building the intelligent image analysis and interpretation unit based on case-based reasoning. In Section 2 we describe the application and the requirements of the system. The case description is explained in Section 3. The initial system architecture is described in Section 4. Finally, we give our conclusions in Section 5.

The application

The following fungal strains (see Figure 1), provided by Perner *et al,* 2002, are used for the study, which were obtained from the fungal stock collection of the Institute of Microbiology, University of Jena, Germany (A) and from culture collection of JenaBios GmbH (B).: Alternaria alternata J 37 (B), Aspergillus Niger i400 (A), Rhizopus stolonifer J 07 (B), Scopulariopsis brevicaulis J 26 (B), Ulocladium botrytis i171(A), Wallemia sebi J 35 (B). A database of images from these species was collected based on fungal stock cultured in the laboratory for the initial development process. All strains were cultured in Petri dishes on 2 % malt-extract agar (Merck) at 24°C in an incubation chamber for at least 14 days. For microscopy fungal spores were scrapped off from the agar surface and placed on a microscopic slide in a drop of lactic acid. Naturally hyaline spores were additional stained with lactophenol cotton blue (Merck).

Image acquisition was conducted using a Zeiss-Axiolab transmission light microscope equipped with a 100x lens and a NIKON Coolpix 4500 digital camera.

Algorithm

1. Collect a sample of the aerial microorganism.
2. Use the equipment given in the section titled "2. The Application" to get a pictorial plate similar to those exhibited in Figure 1 (Fungi and Pollen).
3. Write out the case description similar to Figure 2 (Example of a case Description).
4. Submit the plate and case description to the system whose architecture is given in Figure 3.
5. Continue by taken another sample.

According to Perner *et al,* 2002, the objects in the images are good representatives of the different kinds of fungal spores cultured under optimal conditions and constant climate conditions. However, as it can be seen from the images of *Alternaria alternata* and *Ulocladium botrytis* none of the objects in the image looks like the other ones. There is no clear prototypical object. We can see a high biological variability and besides that we see younger and older representatives of the fungal strains. Depending on the image acquisition conditions we see objects from the side and top view that influences the appearance of the objects. generalization about the objects can not be done at hand rather each case that appears in practice should be stored into the system and the system should learn more generalized description for the different appearance of the same objects over time. All the above explanation suggests that we should take a case-based reasoning approach for the image interpretation (Benyon *et al,* 1999) rather than a generalized approach.

First of all we need to describe the images by image features and then we need to develop a feature extraction procedure, which can automatically extract the features from the images. The features and the feature values extracted from the images together with the name of the fungal spores make up an initial description of the data. We do not know if all image features are necessary. However we extract as much as possible image features from the images, which make sense in some way to ensure that we can mine the right case description from this database. From this initial description of the data we need to identify good representative descriptions for the cases by using case mining methods (Grimnes and Aamodt, 1996).

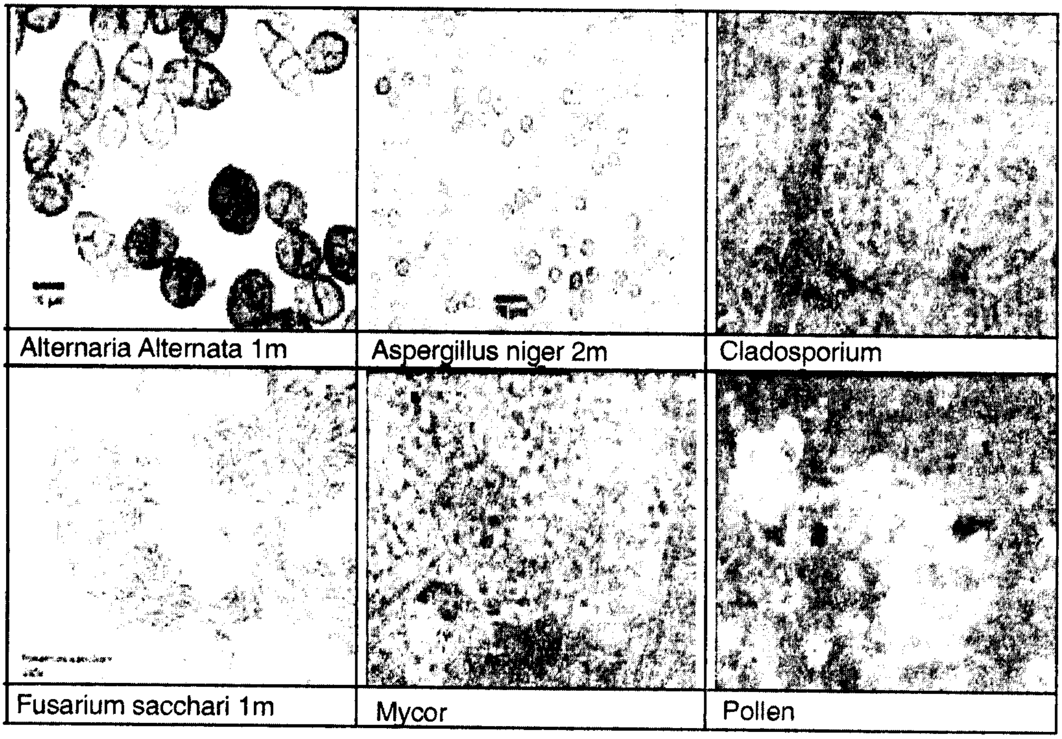

Figure 1: Fungi and pollen.

Case descriptions

An image may be described by the pixel matrix itself or by parts of this matrix (a pixel representation) (Dorge, 2000). It may be described by the objects contained in the image and their features (a feature-based representation) (Perner, 2001). Furthermore, it can be described by a more complex model of the image scene comprising objects and their features as well as the object's spatial relationships (an attributed graph representation (Aha, 2001) or semantic network (Perner, 1999).

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 2). The features are colour, shape, and special properties inside the objects such as structure inside, size, and appearance of the cell contour.

Description
Colour = brown
Object = is structured
Contour = double contour
Shape = bottle-like shapes
Solution Alternaria alternata

Figure 2: Example of a case description.

The architecture

The recent architecture of the system is shown in Figure 3. The cases are image descriptions, which are automatically extracted from the images based on the procedures installed in the feature extraction unit and stored into the case base together with the class name. In a separate image database are kept all images, class names, and image descriptions given by a human operator for later evaluation purposes. The feature extraction unit contains feature extraction procedures that have been identified based on the interview with the expert. We should note here that a particular application requires special feature descriptors. Therefore not all possible feature extraction procedure can be implemented into such a system from scratch. But we hope that we can come up with a special vocabulary and the associated feature extraction procedures for our application on fungi identification.

Similarity between an actual case and cases in case base should be determined based on the Euclidean distance. The initial case base is a flat case base. An index structure and more compact case description are incrementally learnt as soon as new cases are input into the case base. For that we will use decision tree learning and prototype learning methods (Jarmulak, 1998).

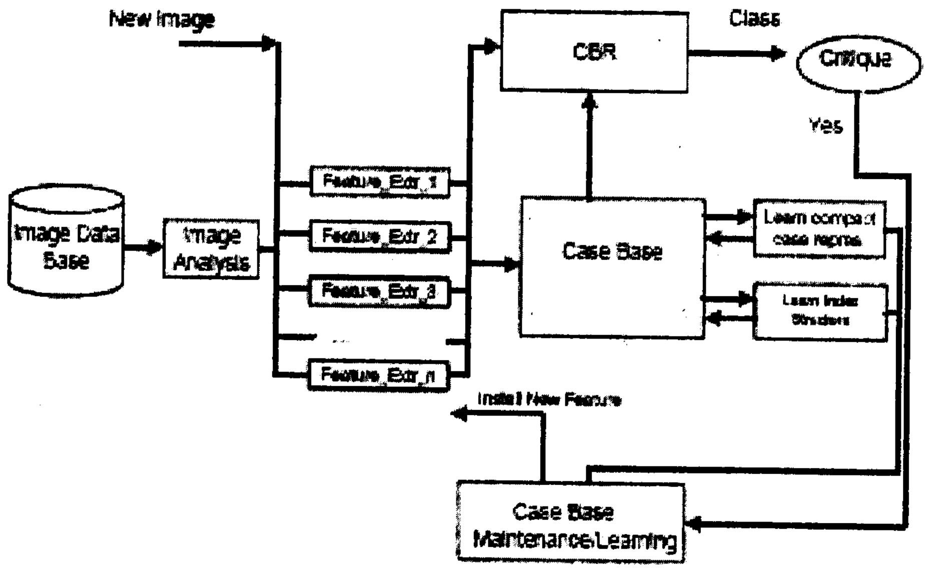

Figure 3: Architecture of the system.

The case base maintenance process will start when the expert criticizes the result of the system. In that case the wrong fungi species has be identified by the system. Then it has to be checked by the system developer whether new features have to be acquired for each case or the case representation should be updated based on the learning procedures. To acquire new features means that the necessary feature extraction procedures have to be developed and for all cases in the case base the new features have to be calculated and inputted into the existing case description. Then the case representation has to be updated as well as the index structure. However, 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 in real life.

Discussion and conclusion

We have described our preliminary study on the development of an image interpretation system for the identification of airborne fungi. Such a system is important for the continuous registration of biological working substances, which have an influence to human health. The paper describes the objects that should be identified from the point of visual inspection. That gives us the basis for the development of the image interpretation system. Features that describe the objects could be identified based on an expert interview and make up the initial case description. Since it is not clear how relevant and discriminative the features are the case base maintenance

process has also to include the identification and installation of new feature extraction procedures. Besides that a representative case description and the index structure over the case base should be learnt incrementally.
The described initial solution is the intelligent heart of the complete system for the automatic registration of airborne fungi. The complete system will comprise a handling unit, an image acquisition unit, and the intelligent image interpretation unit.

Once the airborne fungi have been fully identified, using the system described above, the level of destruction to the human body can be thoroughly studied and explained to the category of people that are prone to the hazard through the mounting of awareness campaigns. The region and period of prevalence are also identified so that immigrants and emigrants into and outside of the region, most especially during the period of prevalence, are adequately catered for through immunization and/or other means of handling such fungi. In third world countries, people do not usually pay attention to any fungi pollutants, more so, if it is not very harmful to the human body and coincidentally many fungi pollutants do not have immediate harming effects on the human body even if they are harmful to man at all.

References

Aamodt A., Sandtorv H. A. and Winnem, O.M. 1998. Combining case-based reasoning and data mining: A way of revealing and reusing RAMS experience, Lydersen, Hansen and Sandorv (Eds.): Safety and Reliability, *Proc. of ESREL'98, Trondheim*, pp. 1345-1351.

Aha I.D. and Watson, I. (Eds.) 2001. *Case-bases reasoning research and developments*, Springer Verlag, lnai 2080, pp. 27-44.

Benyon, F.H.L., Jones A.S., Tovey E.R. and Stone, G. 1999. Differentiation of allergenic fungal spores by image analysis, with application to aerobiological counts, *Aerobiologia 15:* 211-223.

Bergmann R. 1998. Introduction to case-based reasoning. A lecture given at the Centre for Learning Systems and Applications, University of Kaiserslautern.

Dörge T., Carstensen, J.M. and Frisvad, J.C. 2000. Direct identification of pure penicillium species using image analysis, Journal of Microbiological Methods, 41, p.121-133.

Grimnes, M. and Aamodt, A. 1996. A two-layer case-based reasoning architecture for medical image understanding. In: I. Smith and B. Faltings (Eds.) *Advances in case-based reasoning* (pp. 164-178). Berlin: Springer Verlag.

Jarmulak, J. 1998. Case-based classification of ultrasonic B-Scans: Case-base organization and case retrieval. In: B. Smyth and P. Cunningham (Eds.) *Advances in case-based reasoning* (pp. 100-111). Berlin: Springer Verlag.

Perner, P. 1998. Using CBR learning for the low-level and high-level unit of an image interpretation system. In S. Singh (Ed.) *Advances in pattern recognition* (pp. 45-54). Berlin: Springer Verlag.

Perner P. 1999. An Architecture for a CBR image segmentation system, *Journal on Engineering Application in Artificial Intelligence, Engineering Applications of Artificial Intelligence, Vol. 12 (6)*, p. 749-759

Perner P. 2001. Why case-based reasoning is attractive for image interpretation, International Conference on Case-Based Reasoning, ICCBR2001, Vancouver Canada.

Perner,P., Perner, H. and Gunther, T. 2002. Airborne fungi identification by case-based reasoning. A "BIOGEFA" project paper at the Institute of Computer Vision and Applied Computer Science, Arndtstr, 4, 04275 Leipzig.

Perner, P. 2003. *Data mining on multimedia data*, Springer Verlag.