Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bbe

Original Research Article

Application of content-based image analysis to environmental microorganism classification

Chen Li *, *Kimiaki Shirahama*, *Marcin Grzegorzek* *

Research Group for Pattern Recognition, Department ETI, University of Siegen, Siegen, Germany

ARTICLE INFO

Article history:
Received 9 April 2014
Received in revised form
10 July 2014
Accepted 23 July 2014
Available online 19 August 2014

Keywords:
Environmental microorganism classification
Microscopic images
Image segmentation
Shape features
Support vector machine
Late fusion

ABSTRACT

Environmental microorganisms (EMs) are single-celled or multi-cellular microscopic organisms living in the environments. They are crucial to nutrient recycling in ecosystems as they act as decomposers. Occurrence of certain EMs and their species are very informative indicators to evaluate environmental quality. However, the manual recognition of EMs in microbiological laboratories is very time-consuming and expensive. Therefore, in this article an automatic EM classification system based on content-based image analysis (CBIA) techniques is proposed. Our approach starts with image segmentation that determines the region of interest (EM shape). Then, the EM is described by four different shape descriptors, whereas the Internal Structure Histogram (ISH), a new and original shape feature extraction technique introduced in this paper, has turned out to possess the most discriminative properties in this application domain. Afterwards, for each descriptor a support vector machine (SVM) is constructed to distinguish different classes of EMs. At last, results of SVMs trained for all four feature spaces are fused in order to obtain the final classification result. Experimental results certify the effectiveness and practicability of our automatic EM classification system.

1. Introduction

Environmental microorganism (EM) is a microscopic organism living in the natural (rivers, seas, forests, mountains, etc.) and artificial (fields, gardens, fish ponds, aeration tanks, etc.) environments. EM consists of a vast amount of tiny organisms that are invisible to the naked eyes. Just as protozoan and bacteria sized usually between 0.1 and 100 μm, they can only be observed under microscopes. In order to obtain the maximal environmental benefit from different EMs, EM classification is a necessary and important phase in the environmental scientific research process, guiding environmental scientists to get to know different operating conditions [1], habits and characteristics of different EMs [2].

EM classification is also an important indicator for both biological treatment processes and environmental quality evaluations. It works as a natural measurement without any

* *Corresponding authors at*: Research Group for Pattern Recognition, Department ETI, University of Siegen, Hoelderlinstr. 3, D-57076 Siegen, Germany.
E-mail addresses: chen.li@uni-siegen.de (C. Li), kimiaki.shirahama@uni-siegen.de (K. Shirahama), marcin.grzegorzek@uni-siegen.de (M. Grzegorzek).

http://dx.doi.org/10.1016/j.bbe.2014.07.003

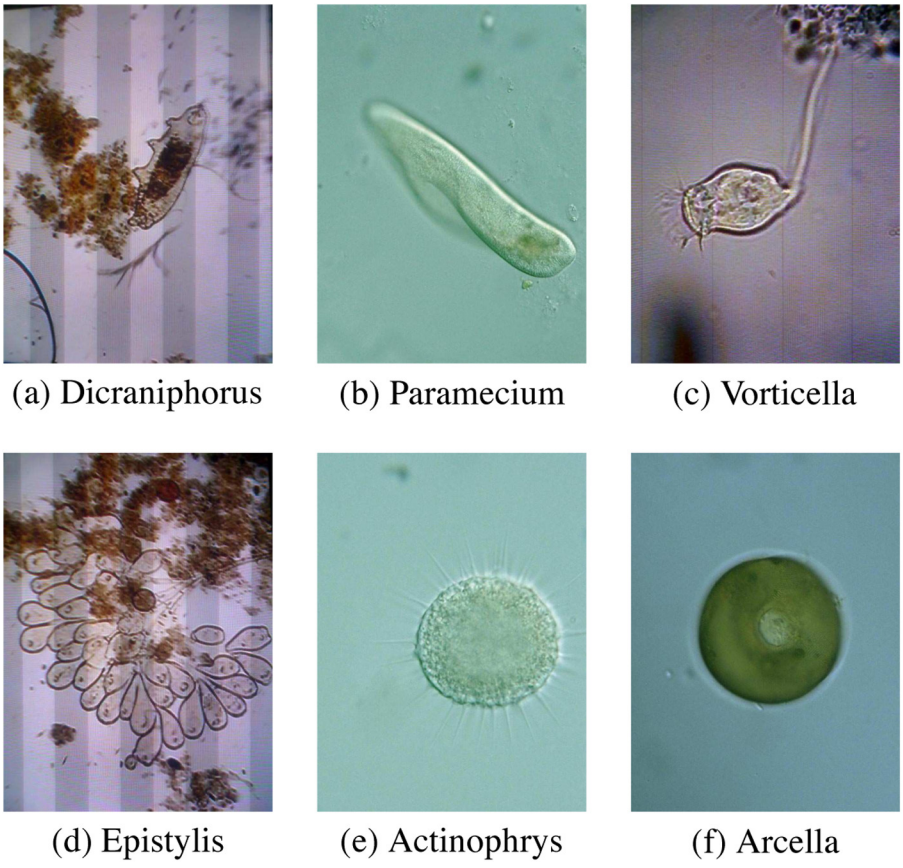

Fig. 1 – Example EMs used as conclusive indicators for the quality evaluation of environmental treatments procedures. (a) and (d) correspond to activated sludge evaluation; (b), (c), (e), and (f) to wastewater treatment evaluation.

additional artificial pollution such as chemical reagent. For instance, the most EMs occurring in wastewater treatment plants are very sensitive to physical, chemical, and operational processes. Fig. 1 shows some examples of EMs. Dicraniphorus and Epistylis (Fig. 1(a) and (d)) are measurements in the process of activated sludge evaluation. Paramecium, Vorticella, Actinophrys, and Arcella (Fig. 1(b), (c), (e), and (f)) indicate the progress in the wastewater treatment process. Therefore, EMs can be considered as conclusive indicators for the quality evaluation of environmental treatments procedures [3,4].

However, it is challenging to distinguish thousands of EMs from each other. There are two traditional ways to identify EMs. The first way is the *morphological method*, in which a microorganism is observed under a microscope and recognised manually based on its shape [5]. This process costs less fund and time, but the training process for the operator is very time-consuming. Furthermore, even very experienced operators are unable to distinguish thousands of EMs without referring to literature. The second way is called *molecular biology method*, which provides a very accurate approach to distinguish EMs by deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) [6,7]. But, the testing processes are time-consuming and the testing equipments are very expensive. In order to overcome problems of these two methods, we have developed a practical and efficient system, in which EM microscopic images are automatically analysed to complete the EM classification task.

Our EM classification system is based on *Content-Based Image Analysis* (CBIA) [8,9]. By CBIA, we mean methods that represent images by numerical feature vectors extracted automatically to represent their perceptual properties (shape, colour, texture, etc.) [10,11]. As shown in Fig. 2, this system simulates the morphological method, where environmental scientists investigate EMs based on their shapes, and detect classes of EMs in different images by a multi-stage analysis approach. Given EM images, the system first conducts image segmentation to obtain shapes of EMs. Then, features which characterise the shape of each EM are extracted from these segmented images. Afterwards, the class of the EM in an image is determined by a classifier based on features extracted from itself. Finally, the result is feedbacked to the user.

Our system has been developed by specifically addressing the issues of (1) the semi-automatic EM segmentation, (2) the robust EM shape description with four different feature extraction techniques, and (3) the EM classification by the late fusion of the outcomes of separate SVMs:

(1) *Semi-automatic EM segmentation:* As the majoring EMs live in very complicated microcosmos full of variegated impurities as in Fig. 1(a) and (d), image segmentation is necessary as a preprocessing phase for obtaining high quality shape information from EM images. For this, we propose a semi-automatic segmentation approach, which refines an EM region that is initially roughly specified by a user.