XIII Congresso Brasileiro de Engenharia Química
XIX Interamerican Congress of Chemical Engineering
I Brazilian Congress of Phase Equilibrium and Fluid Properties for Chemical Process Design

STUDY OF PROTOZOA POPULATION IN WASTEWATER TREATMENT PLANTS BY IMAGE ANALYSIS

M. da Motta[1], M.N. Pons[1], H. Vivier[1], A.L. Amaral[2], E.C. Ferreira[2], N. Roche [1,3], M. Mota[2]
[1] Laboratoire des Sciences du Génie Chimique, CNRS-ENSIC-INPL, 1 rue Grandville, F-54001 Nancy cedex, France
[2] Centro de Engenharia Biológica - IBQF, Universidade do Minho, P-4710 Braga, Portugal
[3] IUT de Marseille, Dept. GCGP, UPRES 884, Université de Marseille, Travers C. Susinni, BP 157, F-13388 Marseille cedex, France

Abstract. *Protozoa are important micro-organisms taking part to the ecosystem balance in wastewater treatment plants. A procedure for their semi-automated identification and counting based on image analysis is proposed. The main difficulty is the segmentation of the protozoa as most of them are in contact with the sludge. The protozoa are characterized by the size of their silhouette (area and length) and three shape factors (elongation, circularity and eccentricity). The identification is performed after projecting the resulting 5D space into a 3D space of Principal Components. The rate of automated identification is actually higher than 50% for some of the species found commonly in activated sludge.*

Keywords: *Image analysis, Protozoa identification, Wastewater treatment*

1. INTRODUCTION

The efficiency of wastewater treatment plants by activated sludge is linked to the bacterial population but also to the protozoa (Nicolau et al., 1997). Different species can be found and have been listed by various authors: Curds and Cockburn (1970a), Martin-Cereceda et al. (1996), Richard (1991), Sasahara et Ogawa (1983), etc. In normal conditions their concentrations are larger than 10^6 protozoa/L. 10^7 protozoa/L corresponds to a very good pollution abatement. On the contrary concentrations lower than 10^5 protozoa /L are indicative of a poor efficiency of the plant (Drakides, 1978). In terms of biomass, protozoa represent between 0.17 and 0.44% of the sludge during the colonization phase but can represent up to 9% at steady-state (Madoni, 1994a). Curds and Cockburn (1970b) have established relationships between the abundance of some species and the sludge loading: they have associated them to the quality of the effluent depending upon the biological oxygen demand (BOD). Table 1 summarizes the predominant groups of protozoa in function of the organic loading. These protozoa have an important role for the good balance of the biological ecosystem: they eliminate the bacteria in excess and stimulate their growth and they promote flocculation (Gerardi et al., 1995). By consuming the free bacteria they help to decrease the effluent turbidity as well as its BOD and its suspended matter content (Curds et al., 1968).

Table 1. Predominant protozoa groups in function of organic loading [from Richard (1991)]

Conditions	Predominant groups
Low organic loading	Stalked ciliates, rotifers and higher invertebrates, especially nematodes.
Optimum organic loading	Good diversity of organisms, dominated by free-swimming and stalked ciliates.
High organic load	Flagellates, amoebae, and small, free-swimming ciliates

Most of the protozoa found in the sludge are ciliated and they can be classified in four main groups: free-swimming, crawling, attached and carnivorous. Table 2 shows that the predominance of one group or another can be an indicator of the efficiency of a wastewater treatment plant by activated sludge. Several authors have applied statistical methods to express the relationships between the protozoa and the operational conditions of the plants. Martin-Cereceda et al. (1996) have used a partial correlation analysis to examine the protozoa of ten wastewater treatment plants at Madrid (Spain) and have established relationships between the protozoa and the plant efficiency (effluent quality and settleability). Using Principal Component Analysis (PCA) with Varimax rotation, Genoveva et al. (1991) have expressed 73% of the process variability by six principal components: the first of these components explains 25% of the variability and takes in account the ciliates.

Table 2. Some relations between protozoa and plant efficiency [from Madoni(1994b)]

Predominant group	Efficiency	Possible cause
Small flagellates	very low	Bad oxygenation of the sludge, too high loading, presence of fermenting substances
Small swimming ciliates (< 50 µm)	low	Contact time too short ; bas oxygenation of the sludge
Large swimming ciliates (> 50 µm)	low	Too high loading
Crawling ciliates	good	
Crawling + attached ciliates	good	
Attached ciliates	decreasing	Unsteady state (discontinuous feeding, sludge wastage)
Small amoebae (with and without flagellum)	very low	Too high loading, not easily biodegradable
Amoebae with shell	good	Low loading, diluted mixed liquor, good nitrification

The protozoa identification and counting needed for the studies previously mentioned has been done manually: this is a very tedious task for an expert. A procedure has been developed by Amaral et al. (1999) for the semi-automated recognition of protozoa by image analysis. The image analysis section, called FlocMorph V0, is embedded into a Visilog™ 5.1 environment (Noésis, Les Ulis, France). The results (size and shape descriptors) are later analyzed by a multivariate method (PCA) for the identification of the protozoa from a database. The procedure was validated on samples regularly taken on a full-scale municipal wastewater treatment plant over a summer period of two months (June and July 1998). However, since that date, other species have been noticed in the samples and the amount of filamentous bacteria has increased drastically, which causes problems in the image treatment. A new version has been developed to take care of the filamentous bacteria and to increase the size of the database.

2. MATERIALS AND METHODS

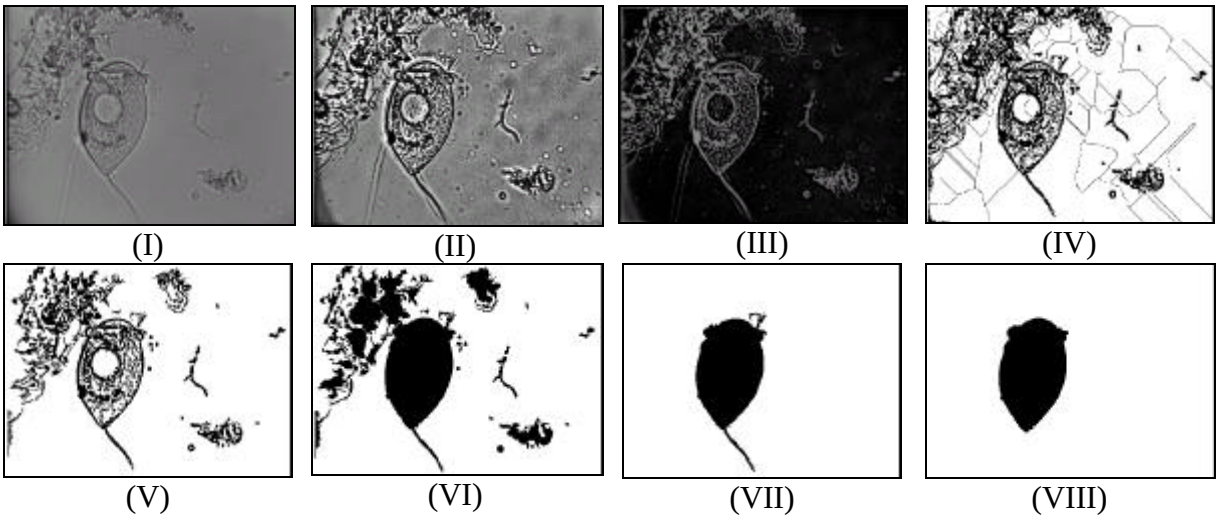
2.1. Sampling and image grabbing

Sludge samples are regularly taken on the wastewater treatment plant of Nancy-Maxéville (350 000 eq. inh.). The delay between the sampling and the image grabbing is about 30 min. The image grabbing system is based on a light microscope (Leitz Dialux 20) and a monochrome video camera (Hitachi CCTV) connected to a PC via a Matrox Meteor board. A mixed liquor drop is deposited on a glass slide and carefully covered with a slip to avoid any mechanical stress on the micro-organisms. For most images a x400 magnification (normal illumination) has been used, except in the case of sets of protozoa (*Opercularia* for

instance) or large rotifers, where a x250 magnification was needed. For each sample 50 images of live protozoa were grabbed by a systematic examination of the slide. The images are stored as TIFF files for subsequent treatment.

2.2. Image treatment

The procedure is called ProtoRec V1 and is implemented in Visilog™5.1 is based on size and shape descriptors obtained on the silhouette of the protozoa. The gray-level image is pre-treated to enhance the contours of the protozoa and is segmented. This is a key step as many protozoa are contact with the flocs and validation by the operator is requested at some points of the procedure. The main steps are presented in Figure 1 and details can be found in Amaral et al. (1999).



(I) Initial image with a x400 magnification (light power = 1V).
(II) Contour enhancement by histogram local equalization (Russ, 1991)
(III) Background suppression by opening (2 iterations) and closing (55 iterations) to remove the halo (Coster and Chermant, 1989).
(IV) Semi-utomated segmentation based on the Euclidian Distance Map (Russ, 1991).
(V) When the protozoan is not in contact with the frame, part of the flocs are eliminated by a border-killing routine. The protozoan contour is closed by openings
(VI) Hole-filling of the silhouette and semi-automated segmentation based on the Euclidian Distance Map.
(VII) Elimination of flocs by a series of erosion and reconstruction of the protozoa silhouette. If flocs are larger than protozoa, they are isolated and discarded by a logical subtraction.
(VIII) Localization of flagella and stalk.

Figure 1. Main steps of ProtoRec V.1

2.3. Measurements

The protozoa are characterized by their size (projected surface, *A*, and length, *L*, given by the maximal Feret diameter, F_{max}) and shape descriptors: elongation, *FS*, circularity, *C* and eccentricity, *E*, calculated from the second-order moments (M_{2x}, M_{2y} and M_{2xy}):

$$FS = F_{max} / F_{min} \quad (1)$$

$$C = P^2 / (4\pi A) \quad (2)$$

where P is the perimeter of the silhouette

$$E = \frac{(4\pi)^2 \left(M_{2x} - M_{2y}\right)^2 + 4M_{2xy}^2}{A^2} \qquad (3)$$

The presence of a flagellum or a stalk is helpful in the identification step but it is not always possible to obtain complete protozoa (with complete flagella or stalk).

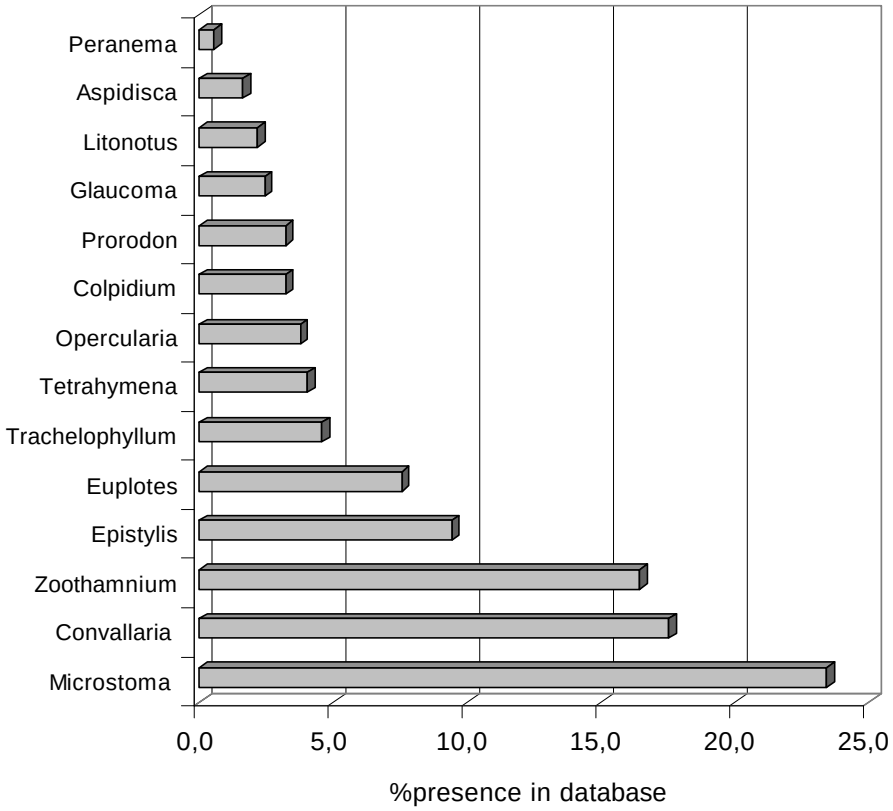


Figure 2. Percentage of presence of the various species of protozoa in the database

Figure 2 gives the percentage of presence of each species in the database. Only protozoa identified clearly by the expert were included in the database. From the total population of protozoa a training set has been defined, with protozoa identified by the expert (Jahn et al., 1979; Madoni, 1994b). A Principal Component Analysis (PCA) (Xlstat™, T. Fahmy, Paris, France) is run on the training data set which contains several individuals of 14 protozoa species, to take into account the variability within each species (Einax et al., 1997).

3. RESULTS

Table 3 gives the eigenvalues obtained from the correlation matrix. The first two principal components, f_1 and f_2, explain 79% of the variability of the training data set. With three components, f_1, f_2 and f_3, 95% of the variability can be explained. No larger improvement is obtained by addition of another component.

The correlation circle (Figure 3) summarizes the relationships between the variables. They are relatively well distributed, indicating that these descriptors can really help to discriminate between the species. As seen in Table 4, *L*, *E* and *C* have a strong effect on f_1, *A* on f_2 and *FS* on f_3.

Table 3. Eigenvalues and degree of explanation of the variability

Factor	f_1	f_2	f_3	f_4	f_5
Eigenvalues	2.4313	1.5397	0.7637	0.1922	0.0731
% variability	0.4863	0.3079	0.1527	0.0384	0.0146
% total variability	0.4863	0.7942	0.9469	0.9854	1.0000

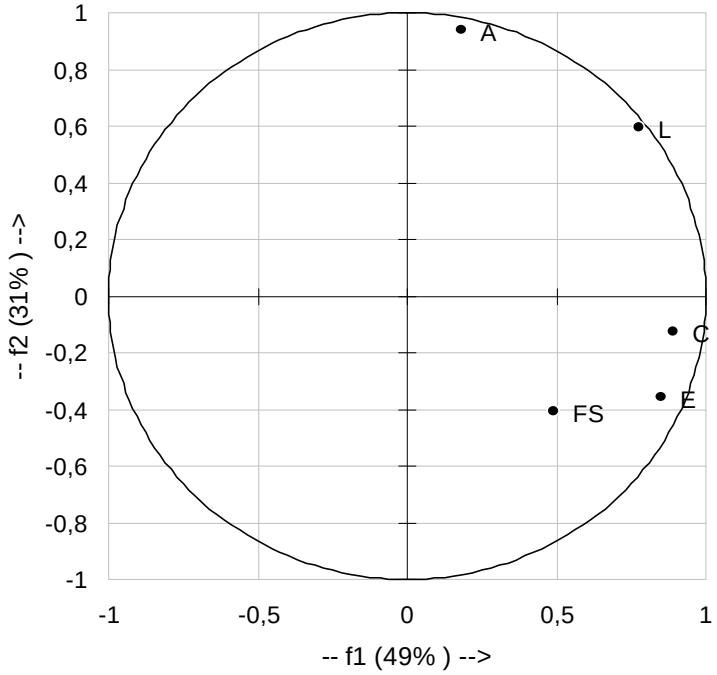


Figure 3. Correlation circle

Table 4. Relationships between the protozoa descriptors and the factors

	f_1	f_2	f_3	f_4	f_5
E	0.8584	-0.3633	-0.1375	0.3348	-0.0155
FS	0.4949	-0.4138	0.7601	-0.0777	0.0007
C	0.8971	-0.1319	-0.3184	-0.2475	-0.1228
A (µm^2)	0.1860	0.9328	0.2548	0.0974	-0.1447
L (µm)	0.7811	0.5908	-0.0255	-0.0576	0.1920

Equations 4 to 6 give the relationships between the co-ordinates in the Principal Component Space (Co_i^j) for each protozoa species *i* along the axis *j*.

$$Co_i^1 = 0.5505\frac{E_i-\mu_E}{\sigma_E} + 0.3174\frac{FS_i-\mu_{FS}}{\sigma_{FS}} + 0.5754\frac{C_i-\mu_C}{\sigma_C} + 0.1193\frac{A_i-\mu_A}{\sigma_A} + 0.5009\frac{L_i-\mu_L}{\sigma_L} \quad (4)$$

$$Co_i^2 = -.2928\frac{E_i-\mu_E}{\sigma_E} - 0.3335\frac{FS_i-\mu_{FS}}{\sigma_{FS}} - 0.1063\frac{C_i-\mu_C}{\sigma_C} + 0.7517\frac{A_i-\mu_A}{\sigma_A} + 0.4761\frac{L_i-\mu_L}{\sigma_L} \quad (5)$$

$$Co_i^3 = -0.1573\frac{E_i-\mu_E}{\sigma_E} + 0.8698\frac{FS_i-\mu_{FS}}{\sigma_{FS}} - 0.3644\frac{C_i-\mu_C}{\sigma_C} + 0.2916\frac{A_i-\mu_A}{\sigma_A} - 0.0292\frac{L_i-\mu_L}{\sigma_L} \quad (6)$$

where m_i is the mean value taken by the parameter i for the whole set of protozoa and s_i the corresponding standard deviation.

In Figure 4 the average position of each species has been plotted in the 3D space of the Principal Components. It can be seen that *V. microstoma* without stalk, *Aspidisca* and

Colpidium are very close one from another. *V. microstoma* can be isolated when its stalk is considered. The same improvement can be obtained for *V. convalaria* and *Opercularia*: the stalk makes the identification easier.

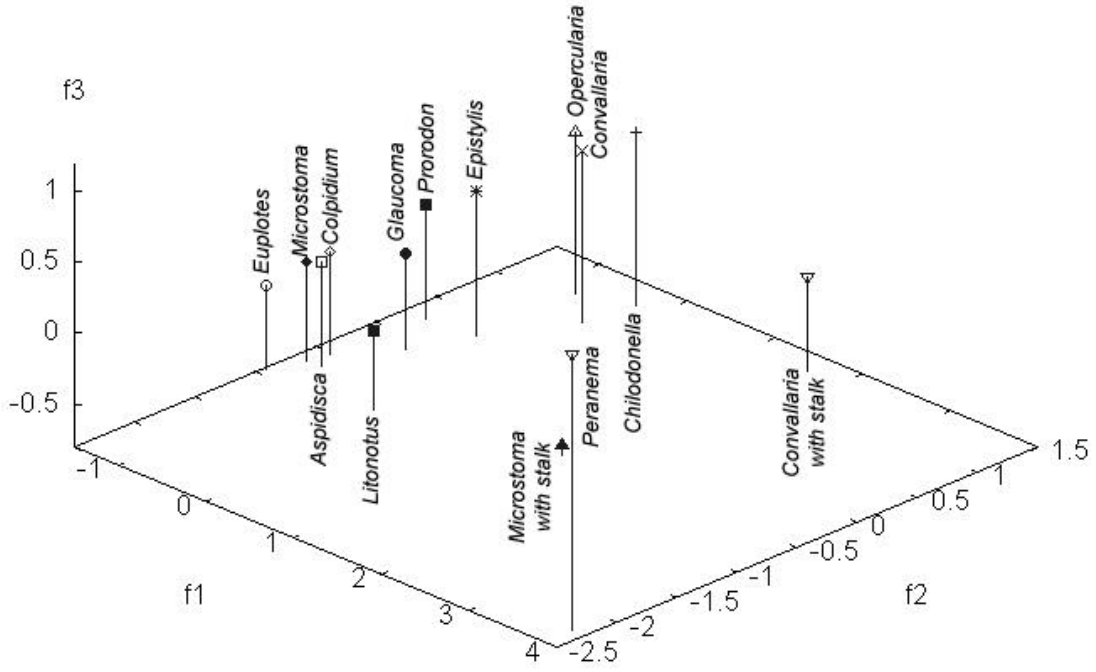


Figure 4. 3D representation of the protozoa species in the PCA space

Table 5. Co-ordinates of each species

	Co-ordinates			Standard deviation		
	f_1	f_2	f_3	f_1	f_2	f_3
Aspidisca	-1.2904	-0.5251	-0.0740	0.3168	0.4242	0.2195
Chilodonella	-0.2118	1.2213	0.4328	0.3624	0.9778	0.5079
Colpidium	-1.1743	-0.6862	-0.0487	0.4035	0.4267	0.2402
Convallaria	-0.2915	0.8309	0.4175	0.4314	0.7395	0.4025
Convallaria with stalk	1.6609	1.2828	-0.1251	**0.9637**	**1.1552**	**0.9238**
Epystilis	-0.7034	0.2515	0.2306	0.6765	0.8059	0.3400
Euplotes	-1.4499	-0.9487	-0.1968	0.3100	0.1895	0.2102
Glaucoma	-0.9144	-0.1706	-0.1048	0.2509	0.1597	0.2389
Litonotus	-0.2705	-0.9075	-0.2256	0.4187	0.3519	0.3911
Microstoma	-1.3307	-0.6904	-0.0877	0.3396	0.2893	0.2560
Microstoma with stalk	1.4275	-0.5740	-0.7155	**1.5784**	**0.6795**	**0.9681**
Opercularia	-0.7166	1.0979	0.3463	0.5036	1.1535	0.3513
Peranema	3.8829	-2.2854	1.1603	**2.1699**	**1.9153**	**3.0906**
Prorodon	-1.2318	0.2190	0.0161	0.3301	0.6311	0.1863
Tetrahymena	-1.4831	-1.3282	-0.0744	0.3781	0.2412	0.2787
Trachelophyllum	-0.6888	-1.6834	-0.1077	0.2714	0.2357	0.5554
Zoothamnium	-0.5211	0.4129	0.3143	0.5645	0.5934	0.3925
Zoothamnium with stalk	0.9939	1.2534	-0.1183	**0.7393**	**0.6112**	**0.4754**

The location of each species and the standard deviation due to the variability within each species are given in Table 5. Flagella and stalks increase the standard deviations as they can have various positions, but they nevertheless improve the identification as the average positions differ considerably when stalk is considered or not. The recognition rate doubles

when the stalk can be taken into account. *Peranema* exhibits very large standard deviations along the three axes due to its small size, its flagellum and its mobility.

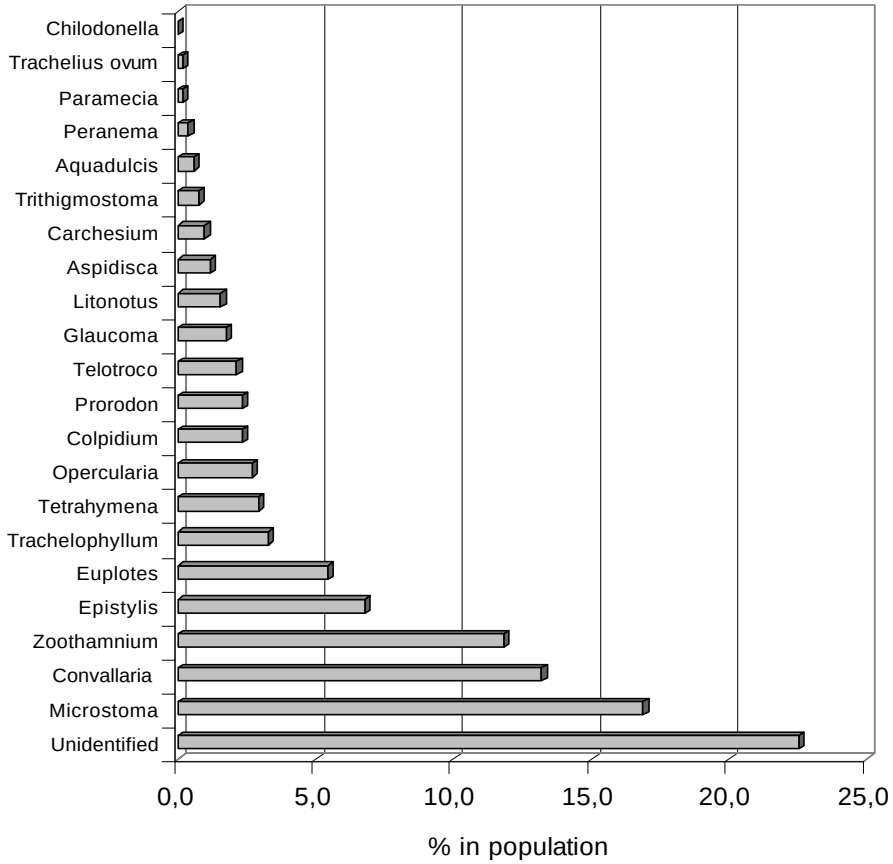


Figure 5. Distribution of protozoa collected over a one-week period.

Figure 5 gives the percentage of presence of the different protozoa imaged during one week and identified by the operator. Some species are not included in the database yet and about 22% of the protozoa could not be clearly identified by the expert. The semi-automated recognition procedure was applied only to the protozoa previously identified by the expert. The protozoa co-ordinates in the PCA space have been computed using equations 4 to 6: the distance of each protozoa to the characteristic position of each species, as given in Table 5, is calculated. The protozoa is assigned to the species for which the distance is minimal. The results obtained by the automated classification have been compared with those found by the operator. Figure 6 gives the rate of successful recognition for the species included in the database.

The rate is larger than 50% for *Zoothamnium*, *Microstoma* and *Convallaria*, that are relatively abundant in the population, as well as for *Trachelophyllum* and *Tetrahymena*. Some species are particularly difficult to recognize: *Peranema*, *Chilodonella* and *Aspidisca* (Figure 7a and b). *Peranema* and *Chilodonella* are new species that have been introduced recently in the database and the limited number of individuals could be a reason for the bad rate of recognition. *Aspidisca* is a small protozoa which is often over the sludge flocs (Figure 7c and d).

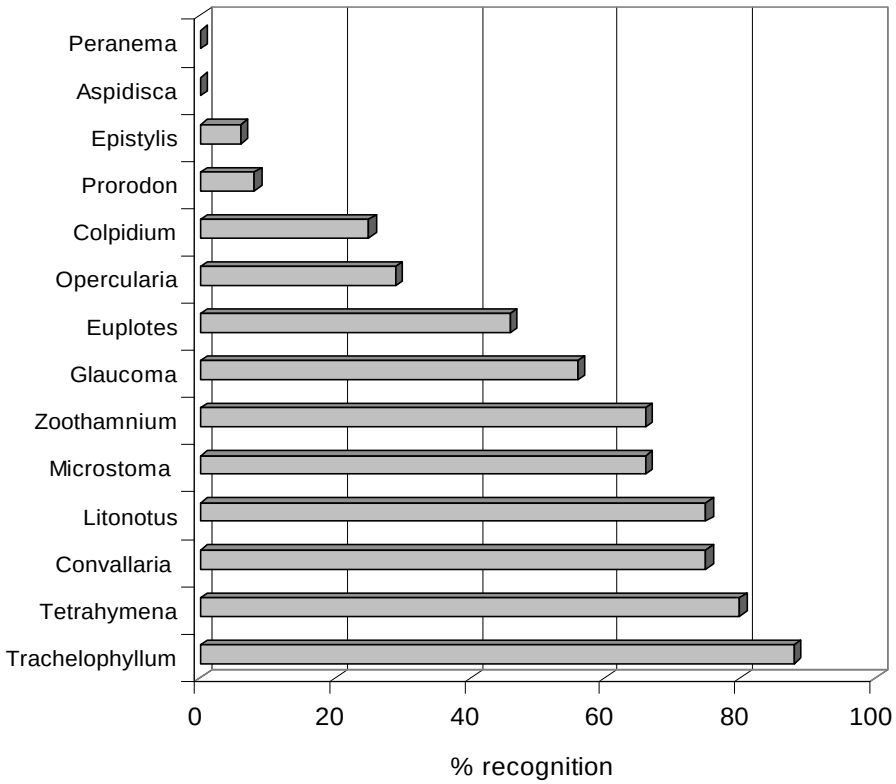
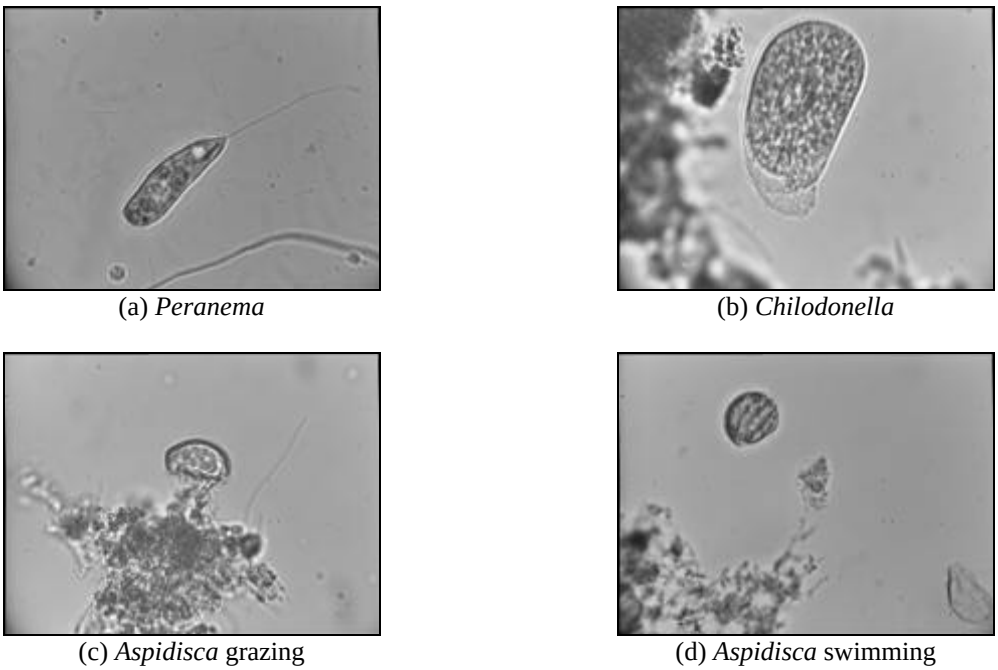


Figure 6. Rate of automated recognition in function of the protozoa species



(a) *Peranema* (b) *Chilodonella*

(c) *Aspidisca* grazing (d) *Aspidisca* swimming

Figure 7. Some protozoa giving difficulties for automated identification

4. CONCLUSIONS

Protozoa are known to be important indicator of the efficiency of wastewater treatment plant. However their manual identification and their counting are tedious tasks. A procedure has been developed to perform these tasks semi-automatically. The segmentation of the protozoa from the sludge flocs is a key step of the image treatment, that cannot be fully

automated at this point. The identification is based on size and shape descriptors of the protozoa silhouette. A database of several individuals belonging to 14 protozoa species has been built. A multivariate analysis of the descriptors is used for the identification of the protozoa.

Although the procedure should be improved, the initial results are promising. Further work is going on to improve the segmentation method of the images as well as the identification by introducting new shape descriptors to characterize the silhouette of the protozoa. In parallel the database is gradually enlarged by addition of new protozoa and introduction of metazoa such as nematodes.

ACKNOWLEDGEMENTS

The authors are thankful to the National Council of Scientific and Technological Development of Brazil (CNPq), the Embassy of France in Portugal and ICCTI.

REFERENCES

Amaral A.L., Baptiste C., Pons M.N., Nicolau A., Lima N., Ferreira E.C., Mota M., Vivier H., Semi-Automated Recognition of Protozoa by Image Analysis, Biotechnology Techniques, 13, 111-118 (1999)

Coster M., Chermant J.L., Précis d'Analyse d'Images, 2e édition, CNRS Editions, Paris (1989).

Curds C.R., Cockburn A., Vandike J.M., An Experimental Study of the Role of the Ciliated Protozoa in the Activated-Sludge Process, Wat. Pollut. Control, 67 , 312-329 (1968)

Curds C.R., Cockburn A., Protozoa in Biological Sewage-Treatment Process – I. A Survey of the Protozoan Fauna of British Percolating Filters and Activated-Sludge Plants, Water Research, 4, 225-236 (1970a).

Curds C.R., Cockburn A., Protozoa in Biological Sewage-Treatment Process – II. Protozoa as Indicator in the Activated-Sludge Process, Water Research, 4, 237-249 (1970b).

Drakides C., L'Observation Microscopique des Boues Activées Appliquée à la Surveillance des Installations d'Epurations : Technique d'étude et Interprétation, TSM-L'EAU, 2, 85-98 (1978).

Einax J.W., Zwanziger H.W., Geiss S., Chemometrics in Environmental Analysis, VCH, Weinheim (1997).

Genoveva E., Tellez C., Bautista L.M., Dynamics of Ciliated Protozoa Communities in Activated-Sludge Process, Wat. Res., 25, No. 8, 967-972 (1991).

Gerardi M.H. et al, Wastewater Biology : The Microlife, A Special Publication, Water Environment Federation, 2nd edition, Alexandria, Virginia (1995)

Jahn T.L., Boove E.C., Jahn F.F., How to Know the Protozoa, The Pictured Key Nature Series, Wm. C. Brown Company Publishers, 2nd edition, Dubuque, Iowa (1979).

Madoni P., Estimates of Ciliated Protozoa Biomass in Activated Sludge and Biofilm, Bioresource Technology, 48 , 254-249 (1994a).

Madoni P., La Microfauna nell'analisi di Qualita' Biologica dei Fanghi Attivi. Manuale di Applicazione, Azienda Gas Acqua Consorziale di Reggio Emilia, Universtità degli Studi di Parma (1994b).

Martin-Cereceda M., Serrano S., Guinea A., A Comparative Study of Ciliated Protozoa Communities in Activated-Sludge Plants, FEMS Microb. Ecol., 21, 267-276 (1996).

Nicolau A., Lima N., Mota N., Madoni P., Os Protozoários como Indicadores da Qualidade das Lamas Activadas, Boletim de Biotecnologia, 56, 14-19 (1997).

Richard M., Activated Sludge Microbiology, The Water Pollution Control Federation, 2e edition, Virginia (1991).

Russ J.C., The Image Processing Handbook, 2nd édition, CRC Press, Boca Raton (1991).

Sasahara T., Ogawa T., Treatment of Brewery Effluent. Part VIII : Protozoa and Metazoa found in the Activated Sludge Process for Brewery Effluent, Monatsschrift für Brauwissenschaft, 11, 443-448 (1983).