FEDERAL UNIVERSITY OF TECHNOLOGY - PARANÁ
GRADUATE PROGRAM IN ELECTRICAL AND COMPUTER ENGINEERING

PHILIPE AMBROZIO DIAS

IN SITU MICROSCOPY FOR ANALYSIS OF FILAMENTOUS BACTERIA: OPTICS AND IMAGE EVALUATION

DISSERTATION

CURITIBA

2016

PHILIPE AMBROZIO DIAS

IN SITU MICROSCOPY FOR ANALYSIS OF FILAMENTOUS BACTERIA: OPTICS AND IMAGE EVALUATION

Dissertation presented to the Graduate Program in Electrical and Computer Engineering of the Federal University of Technology - Paraná as a requirement for obtaining the degree of Master of Science – Concentration Area: Biomedical Engineering.

Supervisor: Prof. Fabio Kurt Schneider, PhD

Co-supervisor: Prof. Hajo Suhr, PhD

CURITIBA

2016

Dados Internacionais de Catalogação na Publicação

D541i 2016

Dias, Philipe Ambrozio
In situ microscopy for analysis of filamentous bacteria : optics and image evaluation / Philipe Ambrozio Dias.-- 2016.
72 f.: il.; 30 cm.

Texto em inglês, com resumo em português.
Dissertação (Mestrado) - Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Engenharia Elétrica e informática Industrial. Área de concentração: Engenharia biomédica, Curitiba, 2016.
Bibliografia: f. 68-72.

1. Águas residuais - Purificação - Processo de lodo ativado. 2. Microscopia. 3. Processamento de imagens - Técnicas digitais. 4. Microorganismos. 5. Bactérias - Monitorização. 6. Métodos de simulação. 7. Engenharia biomédica. 8. Engenharia elétrica - Dissertações. I. Schneider, Fábio Kurt, orient. II. Suhr, Hajo, coorient. III. Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial. IV. Titulo.

CDD: Ed. 22 -- 621.3

Biblioteca Central da UTFPR, Câmpus Curitiba

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
Câmpus Curitiba

Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial

Título da Dissertação Nº. ____

In Situ Microscopy For Analysis Of Filamentous Bacteria:Optics And Image Evaluation

por

Philipe Ambrozio Dias

Orientador: Prof. Dr. Fábio Kurt Schneider
Coorientador: Prof. Dr. Hajo Suhr

Esta dissertação foi apresentada como requisito parcial à obtenção do grau de MESTRE EM CIÊNCIAS – Área de Concentração: **Engenharia Biomédica** do Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial – CPGEI – da Universidade Tecnológica Federal do Paraná – UTFPR, às **13:30h** do dia 29 de fevereiro 2016. O trabalho foi aprovado pela Banca Examinadora, composta pelos professores doutores:

Prof. Dr. Fabio Kurt Schneider
(Presidente – UTFPR)

Prof. Dr. Valdinei Luís Belini
(UFSCAR)

Prof. Dr. Hugo Vieira Neto
(UTFPR)

Visto da coordenação:

Prof. Dr. Emilio Carlos Gomes Wille
(Coordenador do CPGEI)

"A Folha de Aprovação assinada encontra-se na Coordenação do Curso (ou Programa)"

ACKNOWLEDGEMENT

The help of multiple people was essential in my pursuit of a double master's degree. First of all, I would like to thank my family, who have always supported me in my decisions.

My gratitude to Frau Yolanda Mateos and Professor Humberto Gamba, for their help in the submission for the scholarship granted to me (Baden-Württemberg-Stipendium). In addition, I thank Professor Gamba and Professor Zwick, the ones responsible for this partnership between the UTFPR and the Hochschule Mannheim.

I would like to specially thank my supervisors. In Germany, Professor Hajo Suhr, who accepted me in his laboratory and provided me with the necessary knowledge to develop the present project with In Situ Microscopy. Moreover, Prof.Suhr also taught me important aspects of being a researcher and a professor.

In Brazil, Professor Fabio Schneider, who since bachelor studies expertly guided me in both academic and personal aspects. His incentives and advices were/are crucial for my pursuit of an academic career.

In addition, I would like to show my gratitude to Herr Peter Schneider and my colleague Diego Sierra, who helped me during the project and turned the laboratory in an enjoyable environment. My appreciation also extends to my project collegues Thiemo Dunkel and Erika de Leon, who gently came to us at the HS Mannheim to discuss the idea and start our cooperation.

To all my "German" and Brazilian friends, specially Britta, Marcio, Guilherme "Corneta", Thaís, Simona, Michael, Hudson, Larissa, Carla and Raffael. Additionally, to my Brazilian laboratory colleagues Eduardo, Andrea and Charles, for their help with this dissertation and the good atmosphere inside laboratory.

Finally, we acknowledge CAPES and CNPq for the financial support in Brazil.

ABSTRACT

Dias, Philipe Ambrozio. IN SITU MICROSCOPY FOR ANALYSIS OF FILAMENTOUS BACTERIA: OPTICS AND IMAGE EVALUATION. 71 f. Dissertation – Graduate Program in Electrical and Computer Engineering, Federal University of Technology - Paraná. Curitiba, 2016.

In the activated sludge process, problems of foaming and filamentous bulking can occur due to overgrowth of certain filamentous bacteria. Nowadays, these microorganisms are typically monitored by means of light microscopy combined with staining techniques. As drawbacks, these methods are susceptible to human errors, subjectivity and limited by the use of discontinuous microscopy. The present project aims the application of an *in situ* microscope (ISM) for continuous monitoring of filamentous bacteria, providing real-time examination, automated analysis and elimination of sampling, preparation and transport of samples. The ISM previously developed at the Hochschule Mannheim required adaptations for use within wastewater environment, specially in terms of impermeability and development of a cleaning mechanism. With a new objective lens design, the system was simplified to a single tubus and an externally activated cleaning system based on magnetism was created. A proper image processing algorithm was designed for automated recognition and measurement of filamentous objects, allowing real-time evaluation of images without any staining, phase-contrast or dilution techniques. Three main operations are performed: preprocessing and binarization; recognition of filaments using distance-maps and shape descriptors; measurement and display of total extended filament length. A 3D-printed prototype was used for experiments with respect to the new ISM's design, providing images with resolution very close to the ones acquired with the previous microscope. The designed cleaning system has shown to be effective, removing dirt settled above the lens during tests. For evaluation of the image processing algorithm, samples from an industrial activated sludge plant were collected weekly for a period of twelve months and imaged without any prior conditioning, replicating real environment conditions. Experiments have shown that the developed algorithm correctly identifies trends of filament growth rate, which is the most important parameter for decision making. For reference images whose filaments were marked by specialists, the algorithm correctly recognized 72% of the filaments pixels, with a false positive rate of at most 14%. An average execution time of 0.7 second per image was achieved, demonstrating the algorithm suitability for real-time monitoring.

Keywords: *In situ* Microscopy. Filamentous bacteria recognition. Digital image processing. Filamentous bulking and foaming. Filamentous microorganism. Wastewater treatment.

RESUMO

Dias, Philipe Ambrozio. MICROSCOPIA IN SITU PARA ANÁLISE DE BACTÉRIAS FILAMENTOSAS: ÓTICA E PROCESSAMENTO DE IMAGENS. 71 f. Dissertação – Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2016.

Em processos de lodo ativado, problemas de *foaming* e *filamentous bulking* podem ocorrer devido ao crescimento exagerado de bactérias filamentosas. Atualmente, o monitoramento de tais micro-organismos é feito por meio de métodos baseados em microscopia ótica combinada com técnicas de marcadores, os quais apresentam limitações intrínsecas da microscopia descontínua, são subjetivos e suscetíveis a erro humano. O presente projeto visa a aplicação de um microscópio *in situ* (MIS) para monitoramento contínuo de bactérias filamentosas, de forma a possibilitar análise instantânea, computadorizada, sem necessidades de recolher, preparar e transportar amostras. O MIS previamente desenvolvido na Hochschule Mannheim teve que ser adaptado para análise de águas residuais, especialmente em termos de impermeabilidade e da criação de um mecanismo de limpeza. Com a utilização de uma nova objetiva, o novo MIS foi simplificado para um tubo único. Ademais, um sistema de limpeza baseado em magnetismo e ativado externamente foi criado. Um algoritmo de processamento de imagens foi elaborado para reconhecimento e medição de comprimento de estruturas filamentosas, permitindo avaliação em tempo real de imagens sem qualquer técnica de marcadores, contraste de fase ou diluição. O mesmo consiste em três operações principais: pré-processamento e binarização; reconhecimento de filamentos por meio de mapeamento de distâncias e descritores de forma; e, finalmente, medição e visualização do comprimento de cada filamento. Um protótipo construído via impressão 3D foi utilizado para avaliar o novo modelo do microscópio, fornecendo imagens com resolução bastante próxima das adquiridas com a versão anterior do sistema. O mecanismo de limpeza desenvolvido mostrou-se eficaz, capaz de remover partículas sedimentadas sobre a lente durante os testes. Para avaliação do algoritmo de processamento de imagens, amostras de uma planta industrial de lodo ativado foram coletadas semanalmente por um período de doze meses e imageadas sem qualquer condicionamento prévio, replicando condições reais de ambiente. Experimentos demonstraram que o algoritmo desenvolvido identifica corretamente tendências de aumento/decréscimo da concentração de filamentos, o que constitui o principal parâmetro para tomadas de decisão. Para imagens de referência cujos filamentos foram marcados por especialistas, o algoritmo reconheceu corretamente 72% dos pixels atribuídos a filamentos, com uma taxa de falsos positivos de até 14%. Um tempo de execução médio de 0,7 segundo por imagem foi obtido, provando sua aptidão para formar uma ferramenta de monitoramento em tempo real.

Palavras-chave: Microscopia *in situ*. Reconhecimento de bactérias filamentosas. Processamento Digital de Imagens. *Bulking* e *foaming* filamentosos. Microorganismos filamentosos. Tratamento de águas residuais.

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

WWTP	Wastewater Treatment Plant
ISM	*In Situ* Microscope
FISH	Fluorescence *In Situ* Hybridization
TEFL	Total Extended Filament Length
FI	Filament Index
MLSS	Mixed Liquid Suspended Solids
SE	Structuring Element
RRG	Reduced Radius of Gyration
ISMv1	First Generation ISM
CCD	Charge-Coupled Device
PC	Personal Computer
GUI	Graphical User Interface
ISMv2	Second Generation ISM
CAD	Computer Aided Design
SGE	Strongest Geodesic Ends

LIST OF SYMBOLS

$\oplus$	morphological dilation
$\ominus$	morphological erosion
$\circ$	morphological opening
$\bullet$	morphological closing
$\hat{\circ}$	Top-hat transformation of an image
$\hat{\bullet}$	Bottom-hat transformation of an image

TABLE OF CONTENTS

1 INTRODUCTION

Availability of freshwater and energy are essential for human well-being and sustainable socio-economic development. Common in the so called developing countries, crisis of climate (e.g. droughts), poverty, hunger, health and finance are in most cases interconnected by water and energy issues. According to estimates by the WWAP - United Nations World Water Assessment Programme, up to 3.5 billion people in the world do not have proper access to water. Moreover, the global water demand (in terms of water withdrawals) is estimated to increase up to 55% by 2050, as consequence of growing demands from manufacturing (400%), thermal electricity generation (140%) and domestic use (130%). This will progressively reduce freshwater availability, so that more than 40% of the global population is projected to live in regions of severe water stress through 2050 (WWAP - United Nations World Water Assessment Programme, 2014).

To minimize the lack of water, the conservation of water sources and the use of multiple ones are essential. Specially the re-use of wastewater appears as an important strategy, requiring however proper cleaning treatment. Wastewater derives mainly from domestic, commercial and industrial establishments, together with groundwater, surface water and storm water (Economic and Social Commision for Western Asia, 2003). The most relevant contaminants found in wastewater are biodegradable organic compounds, volatile organic compounds, recalcitrant xenobiotics, toxic metals, suspended solids, nutrients (nitrogen and phosphorus), and microbial pathogens and parasites (TCHOBANOGLOUS et al., 2003). To illustrate the importance of proper wastewater treatment, Table 1 resumes the possible impacts of each type of pollutant when released in the environment.

According to DWA (2011), in Germany 8.5 billion m^3 of wastewater were processed by the $5,668$ observed Wastewater Treatment Plants (WWTP) in the year of the publication (part of a total of 9.933 distributed all over the country). In the United States, there are more than $15,000$ wastewater treatment facilities which in total treat approximately 140 million m^3 of wastewater per day (BITTON, 2005). In Brazil, according to IBGE (Instituto Brasileiro de Geografia e Estatística, 2008), more than 8 million m^3 of wastewater are treated per day.

Table 1: Environmental impacts of relevant wastewater components

Contaminants	Environmental impacts
Suspended solids	can lead to development of sludge deposits and anaerobic conditions.
Biodegradable organics	are principally made up of proteins, carbohydrates and fats. Their biological stabilization can deplete natural oxygen resources and cause septic conditions that are detrimental to aquatic species.
Pathogenic organisms	can cause infectious diseases.
Nutrients	nitrogen, phosphorus and carbon are essential nutrients for growth of organisms. When discharged to the aquatic environment they can lead to the growth of undesirable aquatic life.
Priority pollutants	including organic and inorganic compounds, may be highly toxic, carcinogenic, mutagenic or teratogenic.
Refractory organics	tend to resist conventional wastewater treatment. Include surfactants, phenols and agricultural pesticides.
Heavy metals	usually added by commercial and industrial activities, they must be removed for reuse of the wastewater.
Dissolved inorganics	such as calcium, sodium and sulfate are often initially added to domestic water supplies, and may have to be removed for wastewater reuse.

Source: Adapted from (TCHOBANOGLOUS et al., 2003)

The principle of activated sludge is widely used for wastewater treatment (BITTON, 2005). Among the different problems that can arise during these processes are bulking and foam formation, which frequently occur due to the action of filamentous bacteria (SEVIOUR; NIELSEN, 2010). This reveals a necessity of monitoring such microorganisms, which has been typically performed using conventional light microscopy combined with staining techniques. As detailed in further sections, these standard techniques are off-line, time demanding, subjective and susceptible to human errors (KUNST et al., 2000).

An improvement in the examination of activated sludge has been achieved by combining microscopy and image processing, but the use of discontinuous microscopy implies in sampling, transportation and preparation of wastewater samples; a whole process during which information may be altered (DUNKEL et al., 2015).

The *in situ* microscopy appears as a promising technique for an accurate, rapid and simple monitoring of the growth rate of filamentous bacteria. It can be directly

installed in bioreactors or pipelines, eliminating sample collection and preparation of slides. Moreover, it allows an analysis with low statistical error, since large amounts of independent images can be acquired for each sample. In the Mannheim University of Applied Sciences (in german HS Mannheim), an *In Situ* Microscope (ISM) was developed based on pulsed illumination, an objective lens separated from suspension by a quartz window and a CCD-Camera as sensor part (SUHR et al., 1995). This ISM consists of two tubes, one inside each other and with a thin layer of distilled water coupling objective lens and window. Up to now, studies combining this ISM and techniques of image processing have allowed quantification and classification of yeast and animal cells (SUHR et al., 1995; CAMISARD et al., 2002; GUEZ et al., 2004, 2009; WIEDEMANN et al., 2011a, 2011b; BELINI et al., 2013).

Wastewater constitutes a hostile environment for optoelectronic devices, since it contains a considerable amount of different organisms and objects. Moreover, theses structures can possibly settle down above the microscope's window/lens, compromising or precluding the acquisition of quality images. For these reasons, the application of this available ISM for monitoring filamentous bacteria requires some mechanical adaptations, including a device for cleaning the region above the window. In addition, a new objective lens was designed by Professor Dr. Hajo Suhr in collaboration with The Institute for Technical Optics of the University Stuttgart. This structure can be immersed directly in the suspension, so that a simplification of the ISM previous design became possible. Therefore, this work included the conception of a new version of the ISM, simplified to a single tube and containing a cleaning device.

Additionally, the application of this available ISM for monitoring filamentous bacteria requires the development of an image processing algorithm capable of recognition, quantification and classification of filaments. In this way, an online estimation of their total extended length (TEFL) can be achieved and compared to actual standard reference methods. The new ISM and a proper algorithm for image processing were concurrently developed, so that some experiments for evaluation of the algorithm were performed using the preexistent ISM.

This project was developed at both the Laboratory for In Situ Microscopy within the Department of Information Technology in Biotechnology at the Mannheim University of Applied Sciences in Germany and the Laboratory for Image and Signal Processing at the Federal University of Technology – Paraná, Brazil.

1.1 OBJECTIVES

The present project aims at the application of an *in situ* microscope for continuous quantification of filamentous bacteria, providing instantaneous examination of the amount of filamentous bacteria and eliminating sampling, preparation and transport of samples.

1.1.1 SPECIFIC OBJECTIVES

To achieve a proper solution for the proposed problem, the objective of this work can be more specifically divided into the following topics:

- Adjustments in the ISM's design:
  - conception of a new ISM using a new objective lens, which allows simplifying the microscope design to a single tube;
  - construction of a cleaning system to ensure suitable imaging in the presence of so many contaminants.
- Development of a suitable image processing algorithm to allow an automated analysis of filamentous bacteria:
  - recognition of filaments;
  - proper quantification of filaments to allow comparison with standard methods;
  - development of a metric to evaluate how accurate is the detection of filaments provided by the developed algorithm;
  - measurements of total extended filament length.

1.2 STRUCTURE OF THE DISSERTATION

This document is divided into five chapters. Chapter 2 provides the theoretical background for the techniques and strategies applied along this project. Beginning with biological aspects of wastewater environment and filamentous bacteria, also the standard methods for quantification of these bacteria are described and evaluated.

At this point, the concept of *in situ* microscopy is introduced, so that the idea and requirements of this project are justified. In Chapter 3, the materials and methods used for the development of both new ISM's design and image processing algorithm are detailed. Chapter 4 describes the experiments done for validation of the proposed solutions, together with their results and respective discussion. Finally, the conclusions and proposals for continued research are presented in Chapter 5.

2 THEORETICAL BACKGROUND AND PROBLEM DESCRIPTION

2.1 WASTEWATER TREATMENT PROCESS

Wastewater can be defined as any water whose quality has been affected by anthropogenic influence. Its treatment typically consists in the following four steps (BITTON, 2005; TCHOBANOGLOUS et al., 2003):

1. **Preliminary treatment:** operation to remove detritus and coarse materials (e.g. rags, sticks, grit, grease) which could obstruct equipment in the wastewater treatment plant;
2. **Primary treatment:** removal of a portion of solids and organic matter by means of physical processes, such as screening and sedimentation;
3. **Secondary treatment:** chemical (e.g. disinfection) and biological (e.g. activated sludge, oxidation ponds) unit processes are used to treat the wastewater;
4. **Tertiary or advanced treatment:** removal of residual suspended solids, nutrients, pathogens, parasites and sometimes toxic substances.

First established in England in the beginning of the 1900's, activated sludge is a process adopted worldwide as a secondary biological treatment (TCHOBANOGLOUS et al., 2003). A conventional activated sludge system consists in two treatment units, as illustrated in Figure 1:

1. **Aeration basin:** providing aeration by mechanical means, a large number of microorganisms forming the so called "activated sludge" aerobically oxidize organic matter to CO_2, H_2O, NH_4 and new cell biomass (BITTON, 2005). The detention time in the aeration basin varies between 4 and 8 hours;
2. **Solids separation device:** in this second environment the microbial flocs (sludge) formed in the first tank are separated from the treated wastewater by means of gravity sedimentation. A portion of the sludge is recycled back to the aeration basin, while the remainder is wasted to maintain a proper F/M (food to microorganisms ratio)(JENKINS et al., 2003; BITTON, 2005).

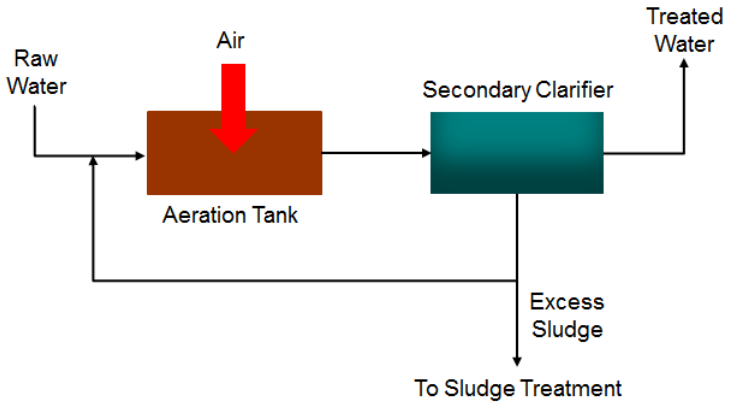


Figure 1: Schematic diagram of an activated sludge process. Raw water enters the aeration tank, where microorganisms forming the so called "activated sludge" degrade organic pollutants into CO_2 and N_2. Thereafter, in the second tank the flocs (sludge) formed in the first tank are separaded from treated water by means of gravity sedimentation. Part of the sludge re-enters the system, while the remainder is wasted.

Source: Adapted from (BITTON, 2005).

Among the main components of activated sludge flocs are bacterial cells as well as other microorganisms, inorganic and organic particles. Bacteria constitute the major component of activated sludge flocs. They are responsible for the oxidation matter and nutrient transformations, besides producing other materials that assist the flocculation of microbial biomass. However, amidst these bacteria are the filamentous ones, which are responsible for some complications in activated sludge processes (BITTON, 2005).

Different problems can occur along the processes of activated sludge and hinder proper sludge settling. Summarized in Table 2, two of these problems are commonly caused by filamentous bacteria: bulking and foam formation. These bacteria are normal elements of the activated sludge, but start to compromise the system when they outcompete the floc-forming bacteria under specific conditions (BITTON, 2005).

A suitable balance between floc-forming and filamentous bacteria results in strong flocs, which settle well in the secondary tank. However, when the filamentous ones are absent or occur in low numbers, small flocs called pin-point are formed, which do not settle well. More critical, filamentous bulking occur when some filamentous organisms are predominant (e.g. *Sphaerotilus, Cyanophyte, Microthrix parvicella*) (SEVIOUR; NIELSEN, 2010), since they bind different flocs and also create some diffuse ones. As consequence, the processes of compaction, settling and thickening are compromised, which in turn leads to formation of a sludge blanket. This blanket

Table 2: Problems in activated sludge wastewater treatments

Problem	Cause	Consequences
Dispersed growth	Microorganisms do not form flocs but are dispersed, forming only small clumps or single cells.	Turbid effluent. No zone settling of sludge.
Slime (jelly), viscous bulking (non filamentous resulting)	Microorganisms are present in large amounts of extracellular slime.	Reduced settling and compaction rates. Virtually no solids separation. In severe cases, overflow of sludge blanket from secondary clarifier.
Pin floc (or pinpoint floc)	Small, compact, weak, roughly spherical flocs are formed; larger aggregates settle faster.	Low sludge settleability and a cloudy, turbid effluent.
Bulking	Filamentous organisms extend from flocs into the bulk solution and interfere with compaction and settling of activated sludge.	High sludge settleability; very clear supernatant.
Rising sludge (blanket rising)	Denitrification in secondary clarifier releases poorly soluble N_2 gas, which attaches to activated sludge flocs and floats them to the secondary clarifier surface.	A scum of activated sludge forms on the surface of the secondary clarifier.
Foaming/scum formation	Caused by nondegradable surfactants, the presence of *Nocardia sp.* and, sometimes, the presence of *Microthrix parvicella*. The last two ones are filamentous bacteria.	Foams float large amounts of activated sludge solids to the surface of treatment units. Then, foam accumulates and putrefies. Solids can overflow into secondary effluent or overflow onto walkways.

Source: Adapted from (JENKINS et al., 1984).

may overflow the secondary clarifier and the process of handling solids can become hydraulically overloaded (JENKINS et al., 2003; KUNST et al., 2000; BITTON, 2005).

The other problem is the formation of foaming by the presence of hydrophobic bacteria like *Nocardioforms*, *Microthrix parvicella* or *Type 1863* (JENKINS et al., 2003). Gas bubbles produced by aeration or metabolism (e.g. N_2) may attach to these foam organisms, which are also connected to flocs. In this way, the flocs become less dense than water and float on the surface, where they accumulate, drain and form a thick brown foam (DUNKEL et al., 2015; JENKINS et al., 2003).

2.2 QUANTIFICATION OF FILAMENTS

Problems caused by filamentous bacteria in activated sludge plants can lead to critical environmental and also economical complications, given that the operating company can become subject to legal and financial penalties. According to Kunst et al. (2000), more than 50% of the problems in wastewater treatment plants are related to bulking sludge, floating scum and foam. This explains the importance of monitoring the behavior of filamentous bacteria. Different methods were developed to observe the growth rate of this type of bacteria in wastewater environments.

Attempts to monitor and control the concentrations of filamentous bacteria have been typically based on optical microscopy (MESQUITA et al., 2011). However, conventional microscopy has already been proven as insufficient for a full characterization of the majority of microorganisms present in wastewater. More recently, some of the most used techniques combine light microscopy with staining techniques or make use of Fluorescence *In Situ* Hybridization (FISH) (JENKINS et al., 2003; KUNST et al., 2000). The latter method targets specific bacteria cells using fluorescent-antibodies, in a way that they can be observed in images acquired with fluorescence microscopes.

In theory, it would be possible to count and measure filaments individually in a given sludge. However, such a task would be not only time consuming, but also could be conducted only if relatively straight filaments are present (EIKELBOOM, 2000). For such reasons, the characterization of filamentous bacteria is typically performed through two methods of quantification: Total Extended Filament Length (TEFL) (SEZGIN et al., 1978) and Filament Index (FI) (EIKELBOOM, 2000; KUNST et al., 2000). The first one, according to Jenkins et al. (2003), consists in counting the number of filaments in a diluted sample transferred onto a microscopic counting chamber, illustrated in Figure 2. They are also classified accordingly to their sizes and the final result is expressed in $\mu m/g$, as showed in Equation 1, where MLSS represents the Mixed Liquid Suspended Solids. Mixed liquor is the content of the aeration tank in an activated sludge system, while MLSS is the total amount of organic and mineral suspended solids in this environment, including microorganisms (BITTON, 2005).

$$\frac{TEFL}{MLSS}[\frac{\mu m}{g}] = \frac{TEFL[\mu m] \text{ in } 1.0ml \text{ of diluted sample x dilution factor}}{\text{MLSS concentration}[\frac{g}{l}]} \tag{1}$$

The second is a simplified method to estimate the amount of filaments

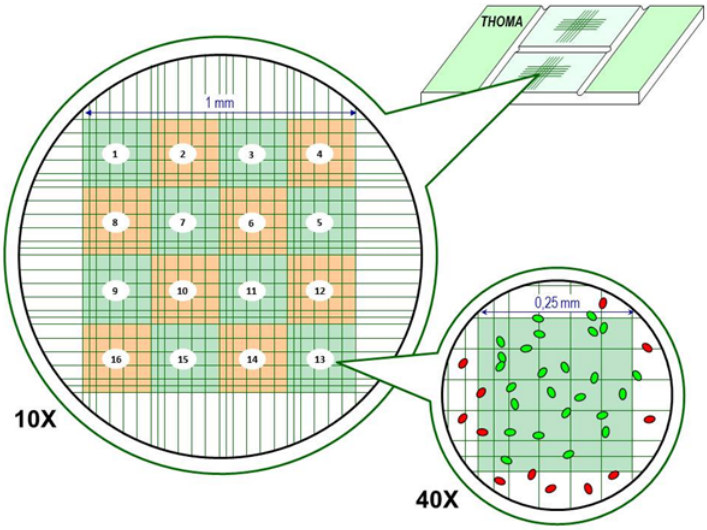


Figure 2: Illustration of a Thoma cell counting chamber.
Source: Extracted from (University of the Basque Country, 2015).

extending from flocs (JENKINS et al., 2003). An index is attributed by comparing the microscopic image of the sludge, at a $100\times$ magnification, with a series of reference photographs of the various FI classes, classifying them in scales from “none” to “very many” filament. One example is the Kunst Index (KUNST et al., 2000), based on a staining technique called crystal violet staining. First, a microscopy slide is prepared with $30\ \mu L$ of sludge and is left to dry. Afterwards, a drop of crystal violet is added and the sample is examined via dark field microscopy ($120\times$ amplification). The reference pictures for the Kunst method are presented in Figure 3.

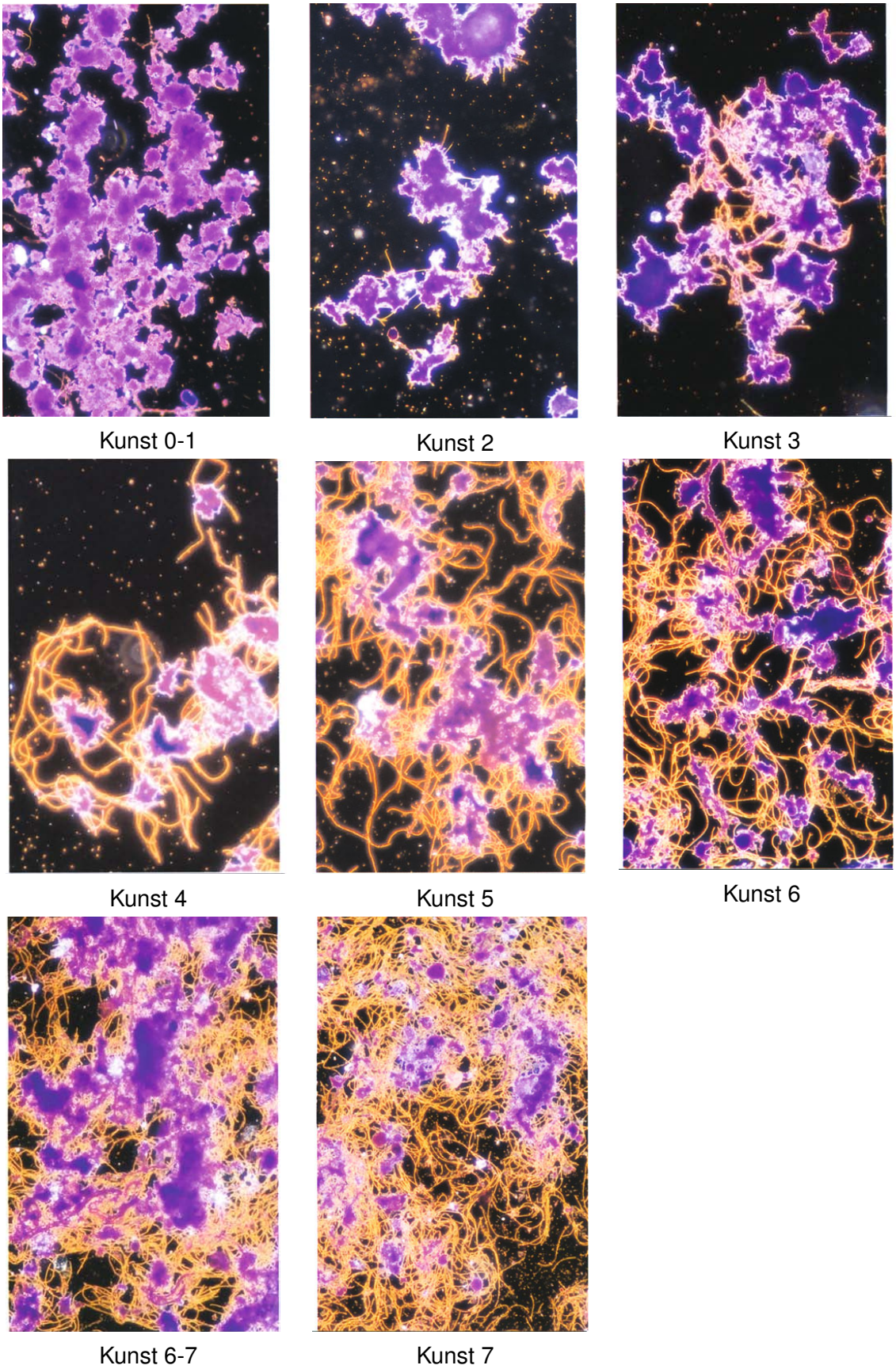


Figure 3: Kunst Index - Reference images for filament counting.

Source: Extracted from (KUNST et al., 2000)

In practice, it has been shown that this method is, in comparison with other ones, quick to carry out and provides very relevant information (KUNST et al., 2000). However, the following points expressed in Eikelboom (2000) highlight the limitations of such methods:

- it is a visual estimation of the amount of filaments. This means that temporal details cannot be examined, as the screening of the slide must be carried out quickly;
- the difference between subsequent class/index can be small, so that the choice of a FI is sometimes arbitrary, subjective;
- in cases of doubt, the FI is rounded up if few flocs are present on the slide and rounded down if many flocs are present;
- if there is an uneven distribution of filaments on the slide under observation, their number is visually averaged;
- typically, observations are made at a $100 - 200$x magnification. Nevertheless, there are cases where higher magnification is necessary (e.g. $300\times$) for visualization of thin filaments and the classification must be revised;
- filaments hidden in the floc cannot be observed and are consequently not included when establishing FI.

2.3 *IN SITU* MICROSCOPY

The quantification of filaments via *ex situ* discontinuous microscopy present different problems and disadvantages. A limited number of images is provided and it is strongly susceptible to subjective interpretations. There is no consensus on the minimum sample size and amount of images required for proper evaluation of activated sludge (KHAN et al., 2015). Furthermore, as summarized in the diagram of Figure 4, they are labor-intensive, requiring sampling, transport, and sample preparation before the evaluation is performed, a whole process during which information may be altered.

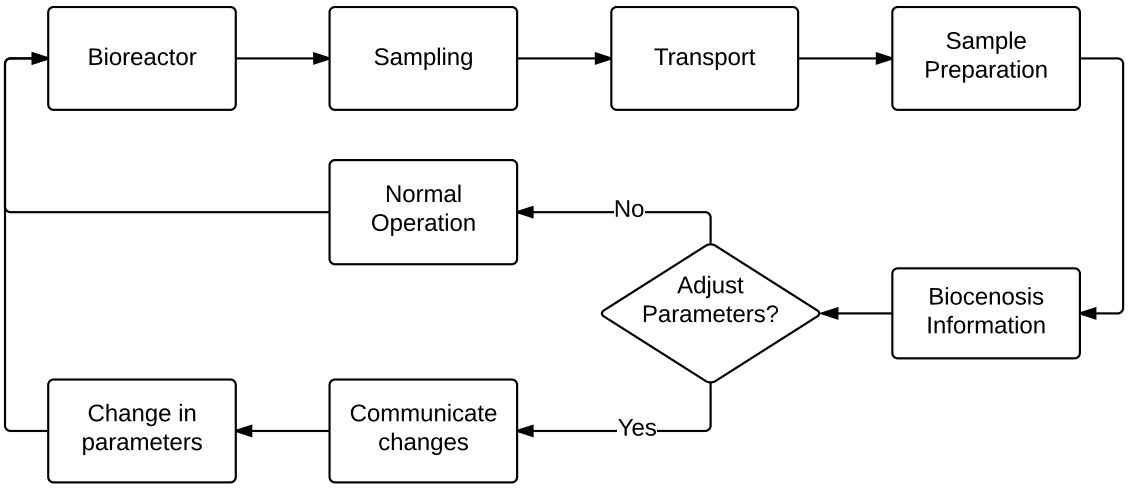


Figure 4: Flowchart of analysis with discontinuous microscopy. First, samples are collected from the bioreactor and transported to the laboratory, where they are further prepared for microscopic analysis. If adjustments in the process of sampling become necessary, changes must be communicated to the personal responsible for sampling.

Source: Adapted from (DUNKEL et al., 2014)

These circumstances reveal the need for a more reliable, consistent method, so that automated image analysis tools have been used for monitoring activated sludge processes. Costa et al. (2013) present an overview of the techniques reported in the literature for quantitative image analysis of wastewater environments. Typically, filaments and aggregates are identified by means of:

- phase-contrast microscopy - images with bright flocs and dark filaments, allowing a classification via direct brightness thresholding (CENENS et al., 2002; JENNÉ et al., 2007);
- fluorescence microscopy - combined with staining techniques, provides images where each type of biological structure (e.g. flocs and filaments) is represented with a specific color (LOPEZ et al., 2005);
- bright-field microscopy - filaments and flocs appear as dark objects against a brighter background, requiring morphological analysis for their distinction (MOTTA et al., 2001; MESQUITA et al., 2010).

Some studies combine different imaging methods for characterization of each structure. In general, bright field microscopy is used for floc characterization, while phase contrast is applied for filaments (JENNÉ et al., 2007; MESQUITA et al., 2010;

AMARAL; FERREIRA, 2005). As disadvantage, these analysis methods still make use of discontinuous microscopy.

At this point emerged the idea of *in situ* microscopy, i.e. the observation of cells inside a culture suspension/bioreactor. Aiming instantaneous, on-line, automatic assessment of concentration of cells and biomass, the concept of *in situ* microscopy was first described by Suhr et al. (1991). With the following advantages (also illustrated in Figure 5), the *in situ* microscopy appears as promising technique for an accurate, rapid and sampling-free monitoring of the abundance of filamentous bacteria under real environment conditions (DUNKEL et al., 2015):

- reproducible, rapid data evaluation with low statistical error due to large amounts of independents (i.e., each image is equivalent to a new slide) images;
- possibility of non-subjective, continuous evaluation (on-line) under real environment conditions;
- the ISM is connected to the bioreactor or pipeline via a standard port, eliminating the necessity of sampling, transport and sample preparation. Breaking of the barrier of sterility and other problems related to sampling are therefore avoided. Besides, the environment is not changed.

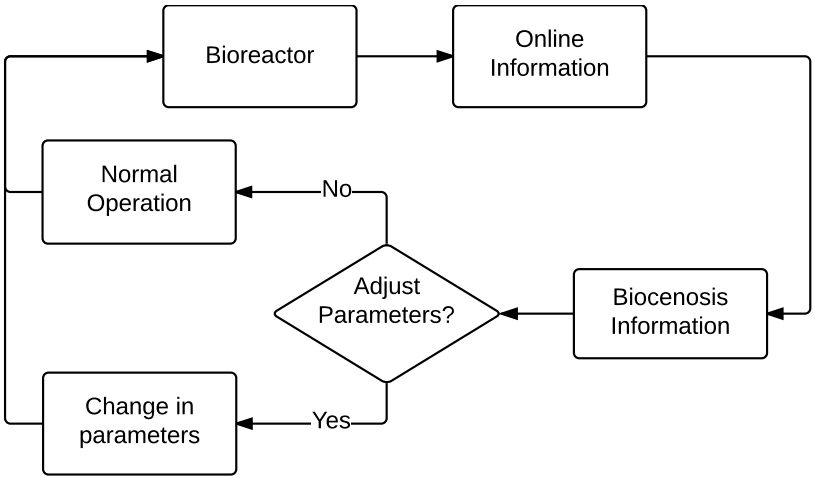


Figure 5: Flowchart of analysis with in-situ microscopy. Since the ISM is directly connected to the bioreactor, on-line information can be obtained without needs of sampling. If adjustments in parameters are necessary, they can be immediately done.

Source: Adapted from (DUNKEL et al., 2014).

The majority of the *in situ* microscopes developed so far are based on bright

field microscopy (BELINI et al., 2013), so that the image processing algorithm must be suitable for real-time identification of filamentous bacteria without any staining, phase-contrast or dilution technique. First, a segmentation between flocs and filaments through direct brightness thresholding is not viable for *in situ* acquired images.

More specifically, none of the imaging analysis systems proposed up to now for characterization of activated sludge flocs are suitable for *in situ* applications (KOIVURANTA et al., 2013). The algorithms proposed by Motta et al. (2001) and Mesquita et al. (2010) are indeed based on bright-field microscopy, but are designed for discontinuous microscopy, where a background image can be acquired and removed during preprocessing and proper dilution can be performed in cases of high concentration. Also the samples analysed with the *in situ* system proposed by Koivuranta et al. (2013) are first diluted and an image from the background is used for image preprocessing.

Thereafter, the application of the ISM for monitoring filamentous bacteria requires mechanical adaptations of the present system for wastewater environment and the development of an image processing algorithm capable of recognition, quantification and classification of filaments. In this way, an online estimation of their total extended length (TEFL) can be achieved and compared to current standard reference methods.

2.4 IMAGE PROCESSING OPERATIONS

Since multiple steps performed in the developed image processing algorithm are based on morphological operations, these procedures are first explained.

2.4.1 MORPHOLOGICAL OPERATIONS

Named after the branch of biology that deals with the form and structure of living beings, mathematical morphology is a branch of image processing which provides tools for representing, describing and analyzing shapes in images (MARQUES, 2011). The *structuring element (SE)*, also known as “window”, is the basic neighborhood structure used in morphological operations, through which geometrical and topological information from an image are extracted. SEs can have arbitrary shapes or commonly used ones, as square, diamond, line and disk.

2.4.1.1 EROSION AND DILATION

The two fundamental morphological operations in image processing are known as *dilation* and *erosion*. The effect of such operations can be well illustrated in binary images. While the first one (denoted by the signal $\oplus$) has the effect of "growing" or "thickening" objects in a binary image, the other (denoted by the signal $\ominus$) has the opposite effect of "shrinking" or "thinning" objects (MARQUES, 2011). The extent and direction of this thickening or shrinking is adjusted by the size and shape of the structuring element. Figure 6 illustrates the effects produced by each operation. The original object shape is given in black, whereas the resulting shape is shown in gray.

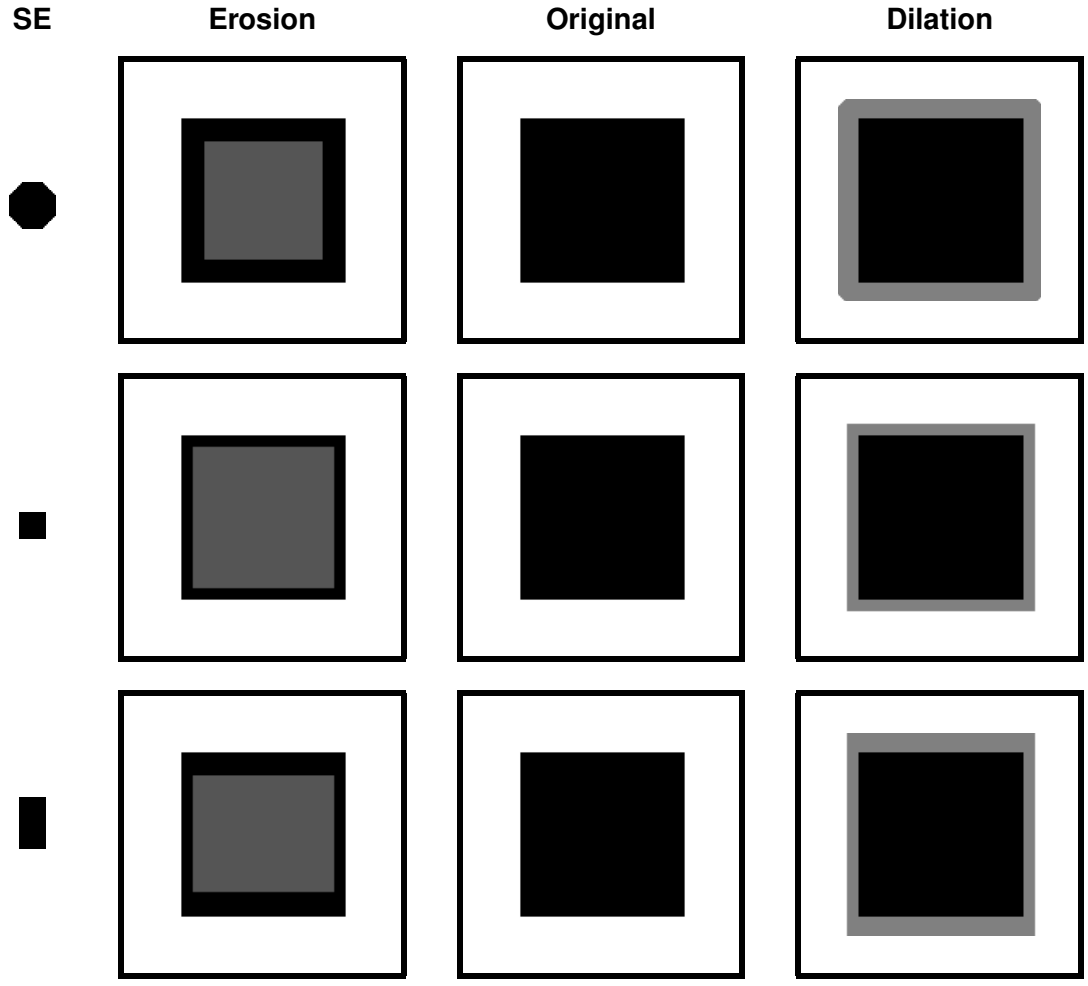


Figure 6: Examples of morphological dilation and erosion. First column presents the structuring elements used in each case. Second column shows the erosion of the image in the third column, while the fourth column shows the respective dilation. The original object shape is given in black, whereas the resulting shape is shown in gray.

2.4.1.2 OPENING AND CLOSING

Two important compound operations, i.e. combinations of dilation and erosion, are known as *opening* and *closing*. Opening operations ($\circ$) are the erosion of an image f by a SE, followed by the dilation of the resulting image by the same SE. Mathematically, this operation is expressed as Equation 2:

$$f \circ SE = (f \ominus SE) \oplus SE. \tag{2}$$

Analogously, closing ($\bullet$) is the dilation of an image f by a SE, followed by the erosion of the resulting image by the same SE. Mathematically, this operation is expressed as Equation 3:

$$f \bullet SE = (f \oplus SE) \ominus SE. \tag{3}$$

As illustrated in Figure 7, the opening operation opens the weak connected inside hole pointed by arrows, while the closing thickly encloses it.

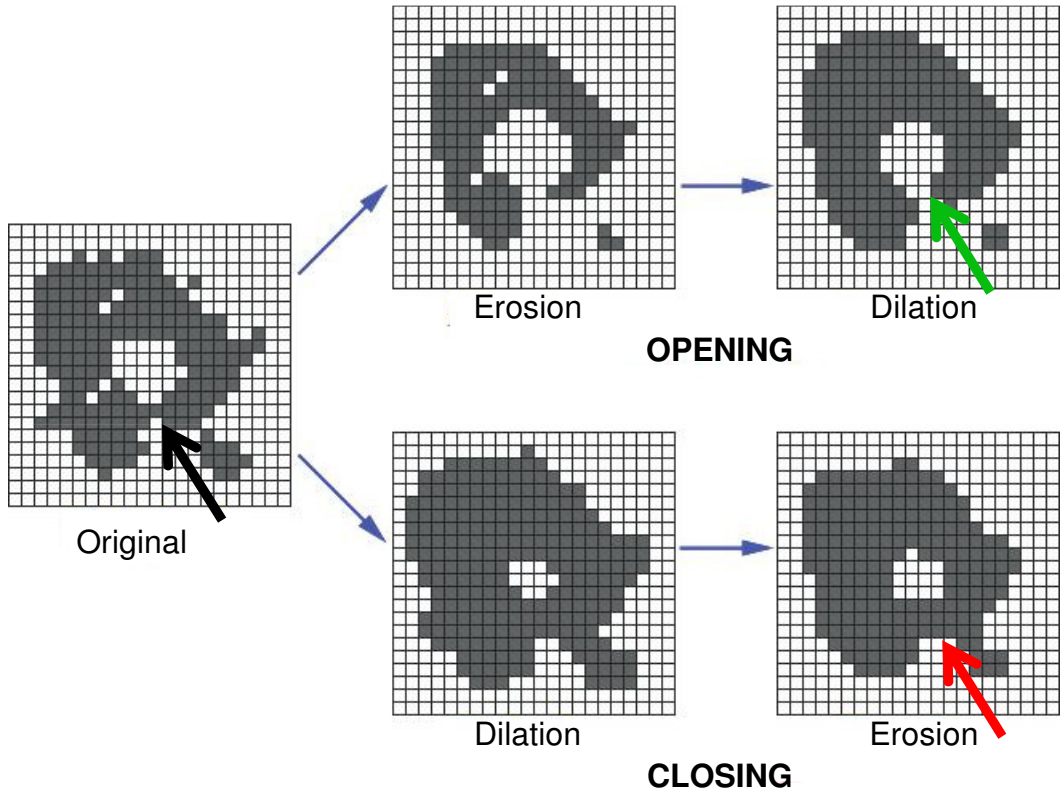


Figure 7: Examples of morphological opening (Top) and closing (Bottom). The effects can be visualized by the opening of the weak connected inside hole (pointed by arrows) in first case, while in the second case the internal hole is thickly enclosed.

Source: Adapted from (RUSS, 2002).

2.4.1.3 OTHER MORPHOLOGICAL OPERATIONS

There are multiple other morphological operations in the literature (RUSS, 2002; GONZALEZ; WOODS, 2007; MARQUES, 2011). Table 3 briefly describes part of them, available in MATLAB through the function *bwmorph*. Some integrate the image processing algorithm discussed on this work, so that their effects are explained and illustrated in Chapter 3 - Section 3.3.

Table 3: Morphological operations - function *bwmorph*

Property	Description
'branchpoints'	Find branch points of skeleton.
'bridge'	Bridges unconnected pixels, i.e. sets 0-valued pixels to 1 if they have two nonzero non-connected neighbors.
'clean'	Removes isolated pixels (individual 1's surrounded by 0's).
'diag'	Diagonal fill eliminates 8-connectivity of the background.
'endpoints'	Finds end points of skeleton.
'fill'	Fills isolated interior pixels (single 0's surrounded by 1's).
'hbreak'	Removes H-connected pixels.
'majority'	Sets a pixel to 1 if five or more pixels in its 3-by-3 neighborhood are 1's; otherwise, it sets the pixel to 0.
'remove'	Removes interior pixels, setting a pixel to 0 if all its 4-connected neighbors are 1.
'shrink'	With n = Inf, shrinks objects without holes to a point. Objects with holes shrink to a connected ring halfway between each hole and the outer boundary.
'skel'	With n = Inf, removes pixels on the boundaries of objects but does not allow objects to break apart.
'spur'	Removes spur pixels.
'thicken'	With n = Inf, thickens objects by adding pixels to the exterior of objects until doing so would result in previously unconnected objects being 8-connected.
'thin'	With n = Inf, thins objects to lines. An object without holes shrinks to a minimally connected stroke, and one with holes shrinks to a connected ring halfway between each hole and the outer boundary.

Source: Adapted from (THE MATHWORKS, INC., 2013).

2.4.2 PREPROCESSING - IMAGE ENHANCEMENT

Typically, the first step of image processing algorithms is related to image enhancement, where operations are performed to improve some key characteristics of an image (e.g. contrast, brightness, edges) and to reduce the presence of noises. There are three major types of image enhancement techniques (SOILLE, 2003):

- point-based techniques: based on the analysis of gray scale values through the whole image, independently of local neighborhoods. Examples are histogram equalization, clipping, thresholding and gamma correction.
- neighborhood-based techniques: operations are performed on local neighborhoods of each image. Representative examples are top-hat and bottom-hat operations.
- transform-based techniques: filtering operations are performed on a transformed version of the image, followed by the inverse transform. Examples are filtering operations performed in frequency domain after Fourier transform.

2.4.2.1 GAMMA CORRECTION

One type of power-law transformation, the gamma correction is an operation typically performed to compensate nonlinearities between equipments used for video/image acquisition and display devices. More specifically, the intensity of light sensed by a camera or reproduced at the display device is a nonlinear function of voltage values, so that a correcting factor is necessary to transform this relationship in roughly linear (MARQUES, 2011). Generally expressed as in Equation 4, gamma correction codes the luminance information into a perceptually uniform space, in such a way that also nonlinear characteristics of the human visual system are corrected:

$$g = k \times f^{\gamma}, \tag{4}$$

where f is the original image, k is a constant, γ is the correction factor and g is the resulting enhanced image.

Gamma-corrected signals also become less sensitive to noise and the operation is helpful for a better encoding of grayscale values, enhancing contrast in certain parts of the grayscale range (GONZALEZ; WOODS, 2007). Figure 8 illustrates the effects of gamma correction with different values of γ.

a) $\gamma = 2$

b) $\gamma = 1$ (original)

c) $\gamma = 0.5$

d) $\gamma = 0.25$

Figure 8: Images illustrating the effect of gamma correction. a) $\gamma = 2$; b) $\gamma = 1$ (original); c) $\gamma = 0.5$; d) $\gamma = 0.25$.

Source: Original photograph by Toni Frissell at Weeki Wachee Springs, Florida, USA, 1947.

2.4.2.2 TOP-HAT AND BOTTOM-HAT OPERATIONS

Frequently used to compensate non-uniform illumination in images, enhancing details in the presence of shading, the *top-hat* transformation ($\hat{\circ}$) is defined as the subtraction of an image f by its opening with a SE, expressed as Equation 5 (GONZALEZ; WOODS, 2007):

$$f \hat{\circ} SE = f - (f \circ SE). \tag{5}$$

Analogously, the *bottom-hat* transformation ($\hat{\bullet}$) is the subtraction of the closing of f with a SE by f, expressed as Equation 6 (GONZALEZ; WOODS, 2007):

$$f \hat{\bullet} SE = (f \bullet SE) - f. \tag{6}$$

As suggested in Marques (2011), they can be combined for purposes of contrast enhancement according to Equation 7, where g is the resulting enhanced image:

$$g = (f + f \hat{\circ} SE) - f \hat{\bullet} SE. \tag{7}$$

2.4.2.3 HIGH-BOOST FILTERING

The application of a high pass filter enhances high-frequency contents but at the expense of the low-frequency contents. Useful for situations where such attenuation of low components is unwanted, high-boost filtering is a technique that enhances high-frequency components without attenuating low-frequency ones (MARQUES, 2011). As expressed in Equation 8, this operation can be performed by multiplying the result of high-pass filtering by a constant and adding it to the original image, acting as sharpening operator specially for edges:

$$g = f + k \times f_{HPF}, \tag{8}$$

where f is the original image, k is a constant value, f_{HPF} is the result of high-pass filtering and g is the image after high-frequency emphasis.

2.4.3 BINARIZATION

One of the main operations in image processing is the *binarization*, a segmentation process by which a number of gray scales in a monochrome image is reduced to only two (black and white). At the expense of information loss, this procedure simplifies and speed up the image's analysis. Segmentation can be done by, for example, thresholding directly the intensity (gray scale) values of an image or a transformed version of the image. One example is the *variance transform*. As expressed in Equation 9, it is defined as the square of the standard deviation. In image processing it represents a statistical measure of gray levels in a square neighborhood,

highlighting regions of transition (GONZALEZ; WOODS, 2007; MARQUES, 2011):

$$Var(X) = \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2, \tag{9}$$

where $Var(X)$ stands for the variance in the region X, x_i refers to one pixel i in the region, σ represents the standard deviation, n is total number of pixels in the region and μ is the average grayscale intensity in the referred region.

2.4.4 DISTANCE TRANSFORMS

The *distance transform* of a binary image is a tool which computes the Euclidean distance transform of a binary image. A number is assigned to each pixel, corresponding to the distance between that pixel and the nearest nonzero pixel of the image (THE MATHWORKS, INC., 2013). This technique is used in this project for measuring thickness of objects, as illustrated within the description of the developed algorithm in Section 3.3.

A particular distance transform is the *geodesic* one, which calculates for each object's pixel the path-length between that given pixel and all the other object's pixels, traversing only the proper object's body as exemplified in Figure 9.

4.83	4.41	4.83
0	3.41	0
0	2.41	3.41
1	0	0
0	0	0

Figure 9: Example of geodesic distance computation. Each pixel receives the geodesic distance between itself and the red pixel, defined as starting point. The white background cannot be traversed during the computation.

The distance between each connected pixel is measured with the concept of *quasi-euclidean distance*, expressed as follows for the distance between an origin

(x_1, y_1) and another point (x_2, y_2) (SOILLE, 2003):

$$\begin{aligned} &|x_1 - x_2| + (\sqrt{2} - 1)\,|y_1 - y_2|\,,\ |x_1 - x_2| > |y_1 - y_2| \\ &(\sqrt{2} - 1)\,|x_1 - x_2| + |y_1 - y_2|\,, \text{otherwise} \end{aligned} \tag{10}$$

2.4.5 OBJECT'S PROPERTIES

To achieve distinction between different objects composing an image, individual properties are relevant. Those information are specially relevant for calculating some shape descriptors, as aspect radio, form factor, roundness, fractal dimension and Reduced Radius of Gyration (RRG). Table 4 lists the information provided by the MATLAB function *regionprops*.

Based on the moments of an object, the RRG is defined as in Equation 11 (PONS; VIVIER, 2000):

$$\begin{aligned} RRG &= \frac{\sqrt{M_{2x} + M_{2y}}}{\frac{D_{eq}}{2}} & D_{eq} &= 2\sqrt{\frac{A}{\pi}} \\ M_{2x} &= \frac{\sum_{i=1}^{N} (x_i - x_g)^2}{N} & M_{2y} &= \frac{\sum_{i=1}^{N} (y_i - y_g)^2}{N} \\ x_g &= \frac{\sum_{i=1}^{N} x_i}{N} & y_g &= \frac{\sum_{i=1}^{N} y_i}{N} \end{aligned} \tag{11}$$

where the position of a pixel *i* of an object is (x_i, y_i), D_{eq} refers to the diameter of a circle with area equivalent to the object's area A, N is the amount of object's pixels and (x_g, y_g) are the coordinates of the object's center of gravity (centroid). The momentum in each dimension is calculated (M_{2x}, M_{2y}) to compute the RRG. For a perfect disc, the RRG is equal to $\sqrt{2}/2$.

Table 4: Object's properties - function *regionprops* (first part)

Property	Description
'Area'	Returns a scalar that specifies the actual number of pixels in the region
'BoundingBox'	Returns the smallest rectangle containing the region
'Centroid'	Returns a 1-by-Q vector that specifies the center of mass of the Q objects composing the region.

Table 4: Object's properties - function *regionprops* (second part)

Property	Description
'ConvexArea'	Returns a p-by-2 matrix that specifies the smallest convex polygon that can contain the region. Only supported for 2-D label matrices
'ConvexImage'	Returns a binary image (logical) that specifies the convex hull, with all pixels within the hull filled in (set to on).
'Eccentricity'	Returns a scalar that specifies the eccentricity of the ellipse that has the same second-moments as the region. The eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length.
'EquivDiameter'	Returns a scalar that specifies the diameter of a circle with the same area as the region.
'EulerNumber'	Returns a scalar that specifies the number of objects in the region minus the number of holes in those objects.
'Extent'	Returns a scalar that specifies the ratio of pixels in the region to pixels in the total bounding box. Computed as the *Area* divided by the area of the bounding box.
'Extrema'	Returns an 8-by-2 matrix [top-left top-right right-top right-bottom bottom-right bottom-left left-bottom left-top] that specifies the extrema points in the region. Each row of the matrix contains the x- and y-coordinates of one of the points.
'Image'	Returns a binary image (logical) of the same size as the bounding box of the region. The on pixels correspond to the region, and all other pixels are off.
'MajorAxisLength'	Returns a scalar that specifies the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the region.
'MinorAxisLength'	Returns a scalar that specifies the length (in pixels) of the minor axis of the ellipse that has the same normalized second central moments as the region.
'Perimeter'	Returns a scalar that specifies the distance around the boundary of the region.
'PixelIdxList'	Returns a p-element vector that contains the linear indices of the pixels in the region.
'PixelList'	Returns a p-by-Q matrix that specifies the locations of pixels in the region.

Source: Adapted from (THE MATHWORKS, INC., 2013).

3 MATERIALS AND METHODS

3.1 ORIGINAL MICROSCOPE DESIGN

First of all, there are by *in situ* microscopy two alternatives for defining the sample volume: the first one is the use of a mechanical device for repetitively enclosing new samples with defined volume in front of the microscope objective lens (BITTNER et al., 1998). The second one, adopted for construction at the HS Mannheim, consists in a purely optical definition of a virtual sample volume by depth of focus (SUHR et al., 1995).

Blurring can be expressed as a monotonic function of the distance between the object and the plane of exact focus, so that if the type and variability of the particles under investigation are known *a priori*, a non-mechanical “virtual” definition of sample volume is possible. In this way, one single two-dimensional image must be enough to evaluate the particle concentration by proper analysis of defocus (SUHR et al., 1995; SCHOLZ et al., 1994). All margins of the virtual sample volume are open and the suspension can flow through the gap between illumination source and quartz window.

After some further developments as reported in Camisard et al. (2002) and Biscaia (2011), the First Genaration ISM (ISMv1) developed at the HS Mannheim can be summarized as follows. Its tubus is schematically shown in Figure 10, while the whole ISM system is depicted in Figure 11 and the process of image formation is illustrated in Figure 12.

It is a pulsed transmitted light microscope, whose illumination is provided by a luminescence diode (DieMOUNT - Germany, λ_{peak} = 640 nm, 2.0 mW at 20 mA), activated by an external circuit described in Biscaia (2011) and guided inside the suspension via optical fiber (WIEDEMANN et al., 2011a). Control pulses $0.5 - 10$ μs wide ensure that, despite the speed of the microorganisms inside a bioreactor (0.1 to 1 m/s), they are still imaged without blurring (BELINI et al., 2013). The fiber-ending (cross section of 1 mm^2) is positioned $\approx$ 0.3 mm above a quartz-glass window separating objective and suspension.

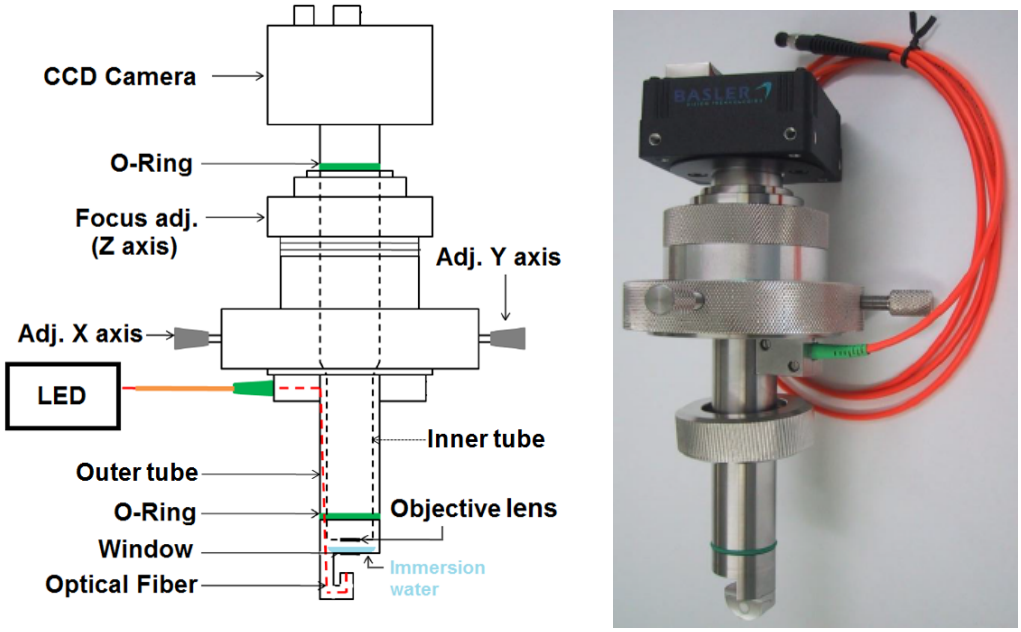


Figure 10: Schematic and photo of the ISMv1 - HS Mannheim. An optical fiber conducts the externally generated pulsed illumination inside the suspension, so that it's endpoint is positioned $0.3\ mm$ above the microscope objective. The adjustment of focus is manually done using the micrometer heads (XY) and the upper-ring (Z), while o-rings assure impermeability.

Source: Schematic extracted from (GUEZ et al., 2004). Photo was taken by the author.

The objective ($\times$40, NA = 0.75) lens is attached to the top of an internal tubus and is optically coupled to the quartz window by means of water immersion. A CCD-Camera (Basler A102f) is attached to the other extremity of the internal tubus. Up to 10 monochromatic images (8-bit, no compression) per second can be acquired, with a resolution of 1392 $\times$ 1040 pixels (pixel size $6.45 \times 6.45\ \mu m^2$) (WIEDEMANN et al., 2011a). A software is responsible for the control of all the system components, triggering both camera and pulse generator according to frequency, gain and brightness defined via an user interface.

The $\times$40 nominal magnification is ensured by the ISM tube length of 160 mm, so that for the given resolution and pixel size an 0.17 $\times$ 0.22 mm^2 object field is captured. The lateral width of the virtual sample volume is directly given by the boundary of the image field (i.e., 0.17 $\times$ 0.22 mm^2), while its third dimension (i.e. its thickness) is defined by depth of focus (SUHR et al., 1995). By other applications, it was estimated to be approximately 10 μm (BELINI et al., 2013). For this work, it is only necessary to ensure that the same virtual sample volume is evaluated by all images. This is ensured by using always the same image processing algorithm parameters.

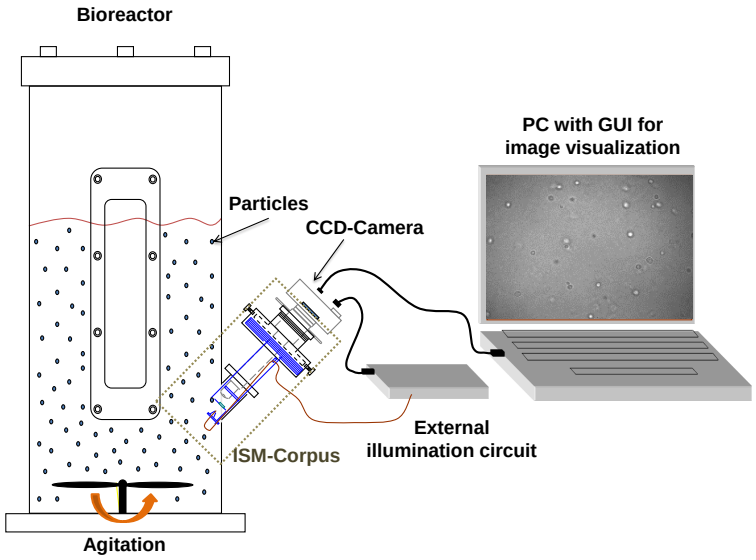


Figure 11: Typical configuration for experiments with the ISM. It can be directly connected to a bioreactor, with it's front-end immersed in the suspension. A constant flow of particles is ensured by means of an agitator. The images formed are transmitted to the PC, where a GUI allows their visualization and adjustment of parameters of image acquisition.

Source: Adapted from (BISCAIA, 2011).

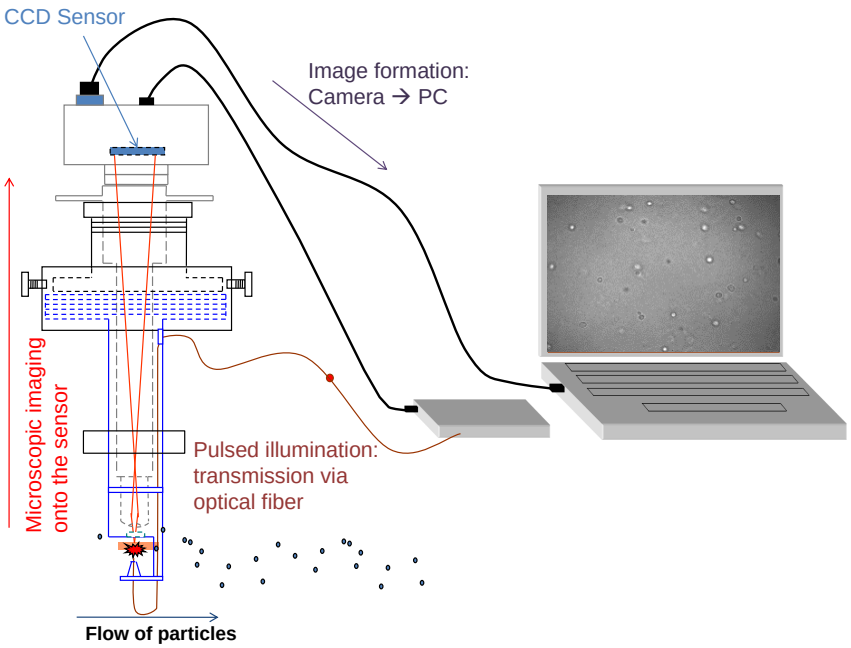


Figure 12: Illustration of ISM's microscopic imaging. Pulsed illumination is generated by an external circuit and guided inside the suspension via an optical fiber, so that it's endpoint positioned closely above the quartz-window illuminates particles passing through this gap. The microscope objective lens, connected to an internal tubus and coupled to the quartz-window by means of destilled water, provides a 40x amplification. The projection reaches the CCD-sensor positioned at the other extremity and the digitalized image is transmitted to the PC.

Source: Adapted from (BISCAIA, 2011).

Research has shown the suitability of this ISM for in-line monitoring average cell volume and concentration of yeast cells (SUHR et al., 1995; CAMISARD et al., 2002), estimation of animal cell-concentration (hybridoma cells) (GUEZ et al., 2004) and cell-viability (GUEZ et al., 2009; WIEDEMANN et al., 2011a, 2011b), besides perspectives for industrial bioethanol production monitoring (BELINI et al., 2013).

3.2 NEW MICROSCOPE DESIGN

3.2.1 SIMPLIFIED CONSTRUCTION

To simplify the two-tube design of the ISMv1, a new objective lens was designed by Professor Dr. Hajo Suhr in collaboration with the Institute for Technical Optics of the University Stuttgart. Composed by three commercial lenses, this objective provides a magnification of 50x instead of the 40x provided by the standard one. It is immersed directly in the suspension, so that neither quartz-window nor immersion water are needed anymore and a new microscope's design with a single tube could be developed. In this way, the microscope requires focus adjustment only a single time, at the installation, so that the ease and safety of operation is much increased. As explained in more details in the following sections, critical points considered in the new design were impermeability, the development of a cleaning mechanism, alignment and distance between the camera and the microscope's window.

3.2.2 CLEANING SYSTEM

Wastewater contains a considerable amount of different organisms and objects, which can possibly settle down between the light source (endpoint of the optic fiber) and the microscope's window/lens. This compromises or even preclude the acquisition of quality images, requiring a device for cleaning the region above the window. Standard solutions as a water pump or other motorized mechanism are not desired for two main reasons. First, such objects could compromise the normal state of the environment under analysis, possibly contaminating it. For an environment like wastewater this is not so critical, but a solution must be universal whenever possible. Second, such solution would also imply in routing cables into the environment, potentially compromising the impermeability of the microscope.

In this project a solution using magnetism was developed. At this point, further

details can not be documented due to a possible future patent application regarding this development.

3.2.3 MICROSCOPE TUBE

Compared to the ISMv1, the new construction (Second Generation ISM - ISMv2) is simpler in that the microscope housing consists of only a single tube. As its previous version, the microscope is designed to be attached to a $25\ mm$ In-gold fitting connection, typically used with bioreactors. For this reason its external diameter has $24.95\ mm$, allowing the microscope to be bolted/screwed to the bioreactor (Figure 13). Close to the sensor head, an o-ring is attached to ensure the hermetic sealing of the bioreactor. In this way the barrier of sterility is preserved, being the outside environment isolated from the inside suspension.

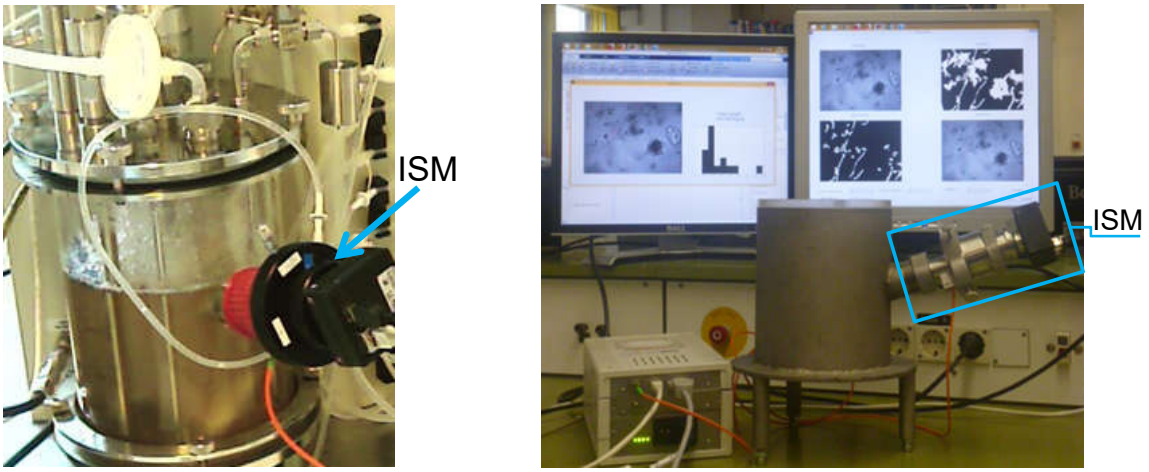


Figure 13: Pictures of the ISMv1 attached to bioreactors, via standard 25 mm In-gold fitting connections.

Source: Adapted from (DUNKEL et al., 2014).

3.2.4 OPTIC FIBER PATH

An external electronic circuit generates the light pulses, which are then guided through the microscope by means of an optic light fiber. A smooth corner-piece leads the fiber inside the near end of the tube. This design is necessary to prevent damages to the fiber due to bending. With the same goal, the fiber is protected by a silicon flexible hose at the tube's head. The endpoint of the fiber is then attached to a ferrule,

filled with epoxy resin and made of stainless steel in order to ensure impermeability and biocompatibility. The distance between the optical fiber's endpoint (light output) and the objective is approximately $0.3\ mm$, so that enough light intensity reaches the lens for proper imaging. Figure 14 shows a simplified CAD diagram of the new ISM's design, while Figure 15 presents a photo of a 3D-printed prototype.

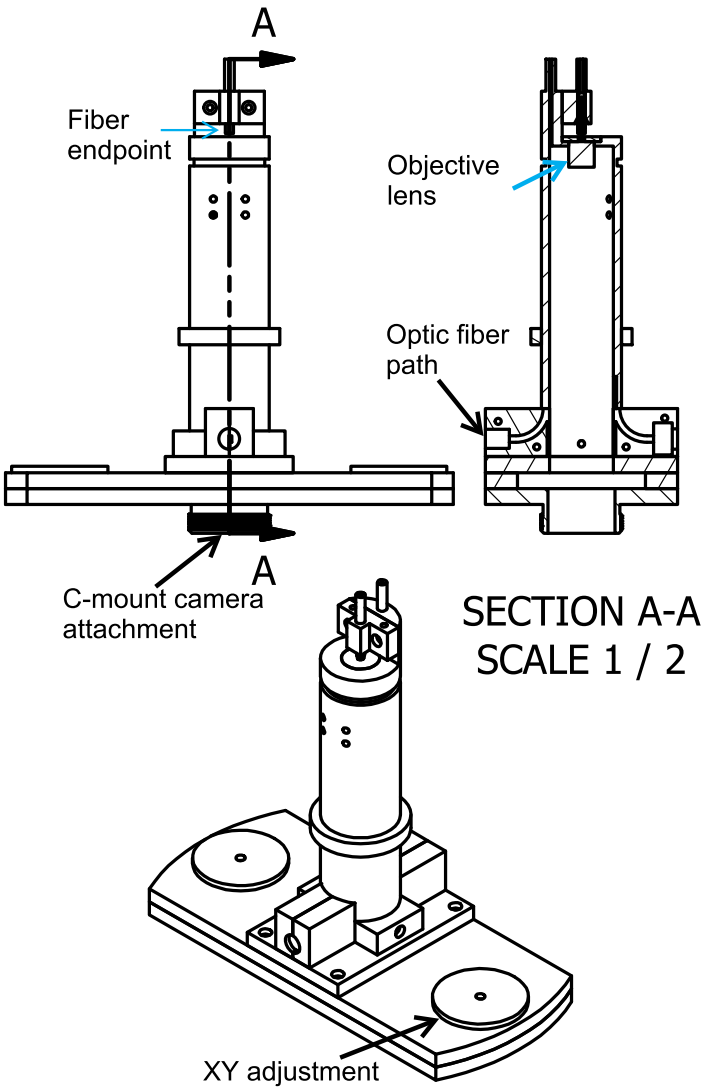
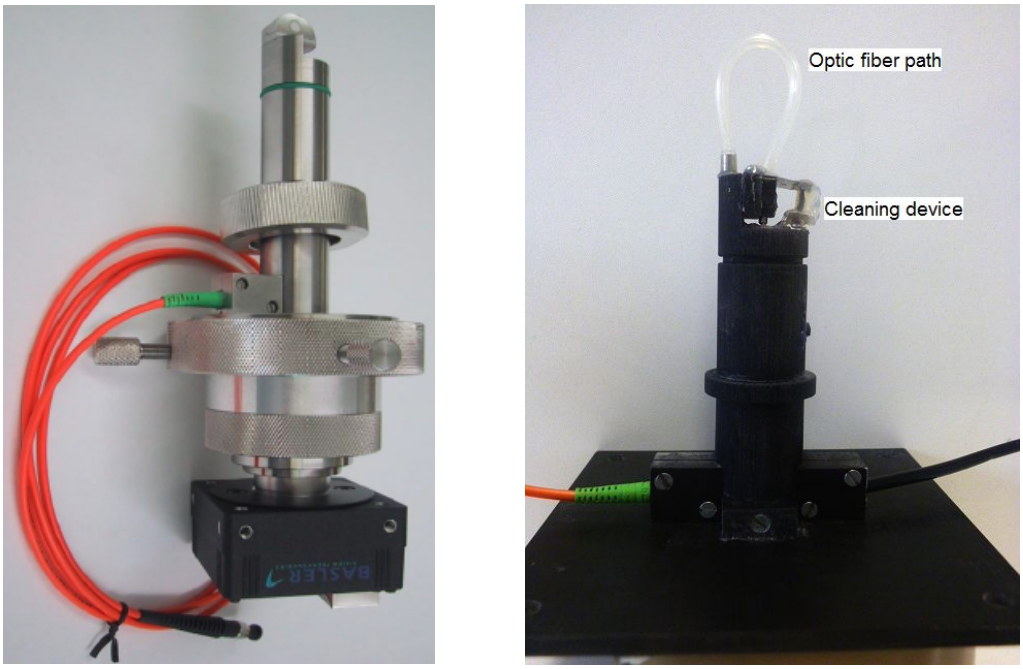


Figure 14: Simplified CAD of the new ISM's design, with front, upper and cross-sectioned views.



(a) Photo of the ISMv1. (b) 3D-Printed prototype of the ISMv2.

Figure 15: Comparison between ISMv1 and ISMv2. At the left side of the ISMv2 is attached the optical fiber. At the top this fiber is encapsuled by a silicon hose, acting as a shield against damages. The cleaning system is controlled by an external button attached to the right side of the ISMv2.

3.3 IMAGE PROCESSING

The image processing algorithm was developed in MATLAB 8.1 (The MathWorks Inc., Natick, MA). One extensively used toolbox is the Image Processing Toolbox, a collection of functions that extend the basics capabilities of MATLAB and enable specialized signal and image processing operations. Furthermore, documentation about its functions and other resources are abundant and well-supplied, on the internet, in its built-in official documentation and in the book *Pratical Image and Video Processing Using MATLAB* (MARQUES, 2011), used as main reference for this project.

Since the new ISM (ISMv2) and the image processing algorithm were developed concurrently, ISM-images acquired with the ISMv1 were used for development and evaluation of the algorithm. ISM-Images taken from wastewater contain not only filaments, but also other agglomerated objects. Since this study focuses on the analysis of filamentous bacteria, an algorithm able to recognize such thin, elongated patterns is necessary. The sequence of operations applied to each image is presented in Figure 16 and detailed in the following topics.

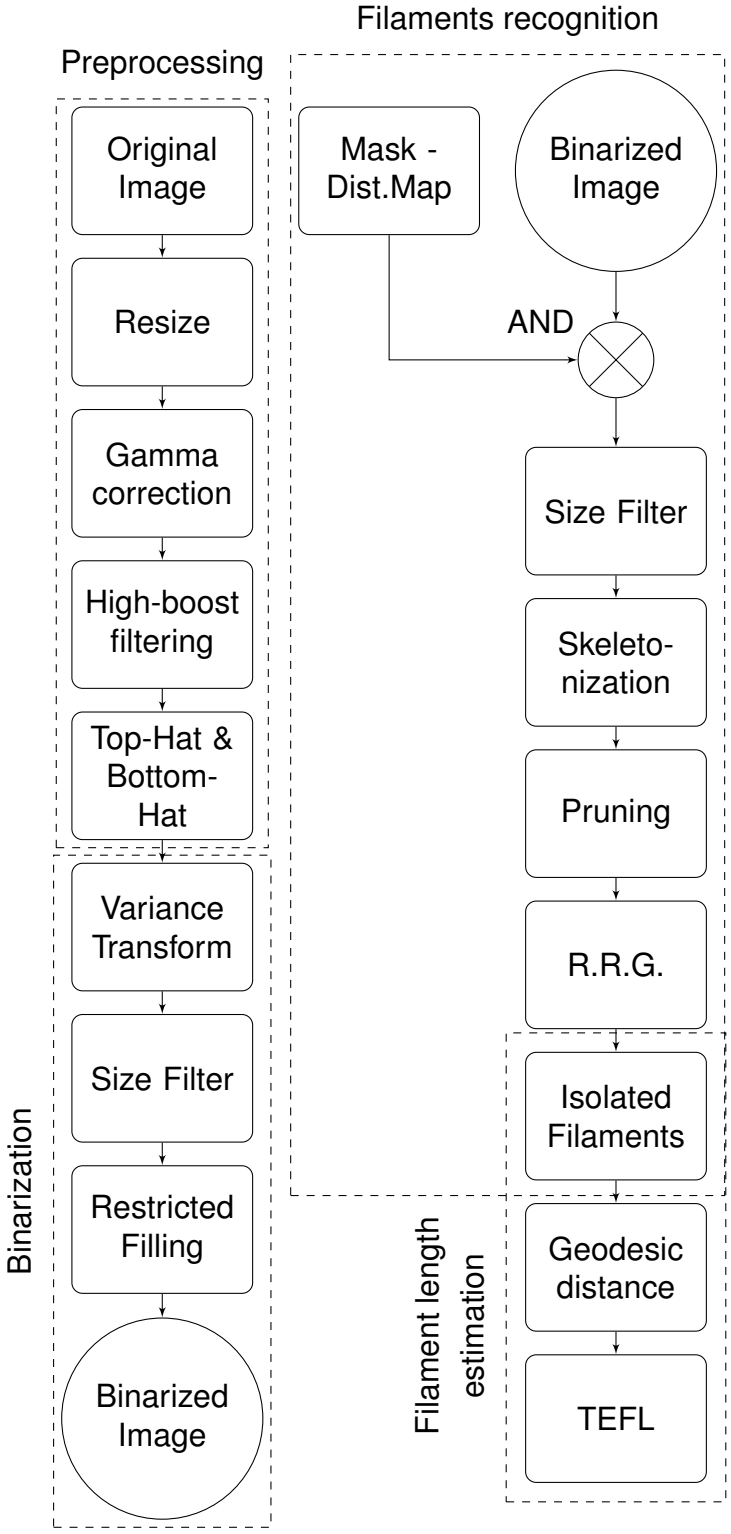


Figure 16: Flowchart of all operations performed for filament detection and length estimation.
Source: Adapted from (DIAS et al., 2016).

3.3.1 PREPROCESSING

In order to save computational time, the image is first resized by a factor of a half in rows and lines (1392 $\times$ 1040 to 696 $\times$ 520 pixels), using bicubic interpolation.

Since the ISM provides resolution of 0.5 μm, even after image reduction the minimal filaments diameter can still be imaged with at least 2.5 pixels, so that there is no major loss of details. Actually, the bicubic interpolation also has an effect equivalent to a low-pass filter, contributing to eliminating some small noises in the image. An average computing time of approximately 0.7 second (around 3.7 times faster than non-resized images) in an Intel Core i7, 3.07GHz, 6GB RAM is achieved. This allows the analysis of thousands of filaments in a short measurement time.

Further, it was verified that an algorithm for Gamma correction according to the Equation 12 provides a better encoding of grayscale values for identification of smaller transitions (RUSS, 2002), as illustrated in Figure 17.

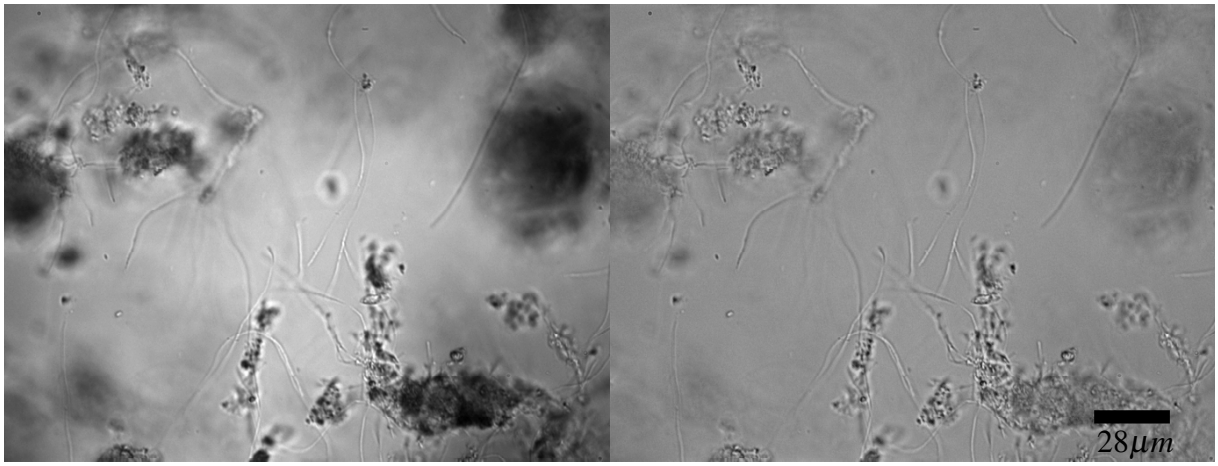
$$g = f^{\gamma} = f^{0.6}, \tag{12}$$

where *f* is the input image and *g* is the output of this step.

Figure 17: Effect of Gamma correction on the encoding of grayscale values. Upper encoding is the original one. Lower bar represents the encoding after gamma correction.

Objects are darker in ISM images if they are in focus or slightly off focus in the direction of the objective. If they are off focus in the other direction, they appear brighter than the background. Moreover, filaments are thin structures whose identification resembles problems of edge detection, so that contrast and high-frequency components in both directions (i.e. darker than background and lighter than background) need to be considered for segmentation.

In the present algorithm, these transitions are enhanced in two stages. First, a high boost sharpening filter is applied, where high-frequency components are emphasized without compromising low-frequency ones. This is performed by adding the original image to its high-pass filtered version. After that, a combination of top-hat and bottom-hat operations (structuring element SE: diamond, radius 3) is applied, expressed as in Equation 7. The top-hat operation ($\hat{o}$) enhances sharp bright peaks, while the bottom-hat operation sharpens dark valleys in the image. Figure 18 illustrates the effect of this step.



(a) Original image. (b) Enhanced image.

Figure 18: Effect of high-frequency and contrast enhancement. Objects in the image are highlighted and can be more easily separated from the background.

3.3.2 BINARIZATION

After image enhancement, the segmentation between objects and background is performed using the concept of variance transform, calculated in 3x3 windows for the original gray values. Regions of transition are highlighted, so that the virtual sample volume can be ultimately defined by variance thresholding with a fixed, optimal threshold value. Objects outside this range (i.e. out-of-focus and poorly focused objects) are imaged with so much blur that they do not pass this threshold and are therefore discarded. The resulting binary image contains the objects in focus present in the original image and some small debris, which can be removed by means of size filter.

According to reference data in Dunkel et al. (2015), the most common species of filamentous bacteria present in the activated sludge under investigation (e.g. *M.parvi-cella*, *Type 1863*, *Type 0092*, *Type 0914*, *G.amarae*) have diameters of $0.8 - 1.2\ \mu m$ and lengths of $10 - 200\ \mu m$ (JENKINS et al., 2003). Considering the ISM's $40\times$ magnification, the image resizing by a factor of a half and its pixel length of $6.45\ \mu m$, the smallest filament diameter corresponds to 2.5 pixels in the final image:

$$\textit{diameter}\text{: } 0.8\ \mu m \xrightarrow{\times 40} 32\ \mu m \xrightarrow{\div 2 \times 6.45\ \mu m} \approx 2.5\ \textit{pixels}$$

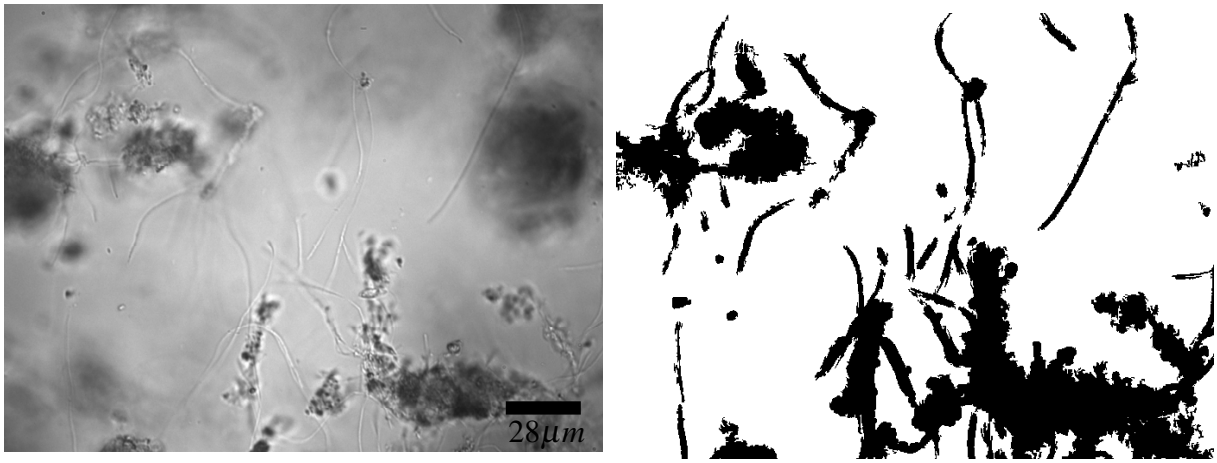
For a rough estimation of area, filaments can be approximated as a rectangle whose area corresponds to width multiplied by length. Considering the smallest width

and length expected, an analagous calculation indicates that true filaments must have areas larger than 77.5 pixels:

$$\textit{length: } 10\mu m \xrightarrow{\times 40} 400\mu m \xrightarrow{\div 2\times 6.45\mu m} \approx 31 \textit{ pixels}$$

$$\textit{area: diameter} \times \textit{ length} = 31 \times 2.5 \approx 77.5 \textit{ pixels}$$

Thereafter, all objects with area smaller than 77 pixels are removed. Moreover, after binarization both filaments and agglomerates contain some small holes not present in the original image, so that these structures need to be filled for better approximation of their original forms. However, filaments can be curved and as consequence form enclosed regions, so that a process of restricted filling is applied: only holes with diameter smaller than the ones expected from true filaments (i.e., 2.5 pixels) are filled. Finally, a binarized image as the one shown in Figure 19 is achieved.



(a) Original image. (b) Binarized image.

Figure 19: Binarized ISM-image of filamentous bacteria in activated sludge. Objects composing the image are separated from the background.

3.3.3 RECOGNITION OF FILAMENTS

In the absence of specific staining or phase-contrast differences, thickness features are used for distinguishing between filaments and other objects. For thickness determination, the *distance transform* is applied. A number is assigned to each black pixel of the binary image, corresponding to the euclidean distance between that pixel and the nearest white pixel composing the background, as illustrated in Figure 23b.

A proper distance thresholding provides a binary image containing only agglomerates, somewhat eroded since their borders are also removed together with

filaments and other small objects. After a suitable dilation by adding pixels to the exterior of objects, their original form is reconstructed and a binary mask containing only flocs is obtained. By means of simple binary *AND* operation, the floc regions are removed from the original binary image (Figure 23a). Followed by the reaplication of size filter, this step results in an image with only thin components as illustrated in Figure 23c.

3.3.3.1 SKELETONIZATION

At this point, only filaments and some fragments, typically parts from flocs, remains in the image. The two groups can be differentiated through measurement of elongation, since the filamentous bacteria present in wastewater are mostly elongated objects which do not exhibit many directional changes. Yet, a proper evaluation of such characteristic shall focus on the object's spine, i.e. the medial axis along its skeletonized version. For this reason, first a skeletonization is performed by successively thinning the objects, removing boundary pixels without allowing connected parts to break apart. Figure 23d shows the output of this step.

3.3.3.2 PRUNING

The skeletonization process yields objects spines together with their branches, which are unwanted for proper length and elongation measurements. For their isolation, we define spines as the shortest path between the two most distant endpoints of a skeleton, as suggested by Soille (2003). Initially, each object is individually recovered in the smallest rectangular frames containing them.

The procedure for elimination of extremities (pruning) is based on the concept of *geodesic distance transform* (SOILLE, 2003). Also known as the Strongest Geodesic Ends (SGE), the two most distant endpoints are recognized by finding the maximal geodesic distance (i.e., the longest path) along the skeleton when starting from each of the four candidate points. To reduce the computational cost, the four endpoints nearest to the four borders of the frame are preselected as possible candidates, as shown in Figure 20a. If two or more endpoints touch a border, the one with smaller pixel index is chosen, i.e. the one closest to the upper-left corner.

Once the SGE are identified, the actual spine can be discriminated from its spurious branches by determining the shortest path connecting the SGE. The

technique to identify the spine pixels is related to an algorithm attributed to Carl F. Gauss for calculating the sum of an arithmetic progression $1+2+3+...+n$. To every pixel is assigned the sum of two numbers. The first summand is the length of the path to one of the endpoints (known as geodesic distance function), while the second summand is the path-length to the other endpoint. Figure 20b illustrates the distance values obtained using one endpoint as origin, while Figure 20c illustrates it for the other endpoint.

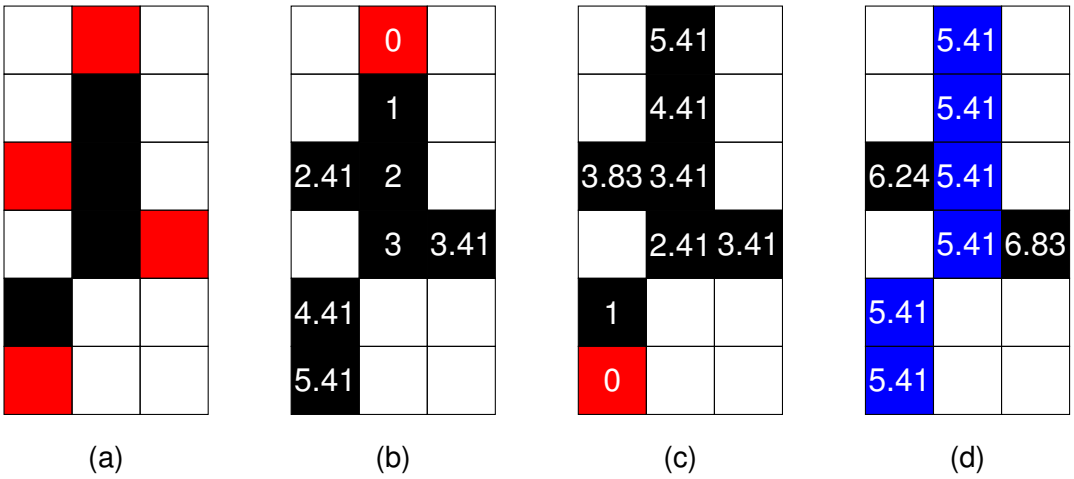


Figure 20: Spine recognition using geodesic distance transform. a) Red pixels illustrate the four endpoints verified for computation of spines; b) and c) Geodesic distance to each skeleton's pixel, starting from endpoints marked in red; d) after summing the results of b) and c), the spine (in blue) can be identified by local regional minimum.

Source: Extracted from (DIAS et al., 2016).

As result, all pixels belonging to the spine will carry the same sum-value which is precisely the geodesic path length between the two endpoints. Pixels outside the spine present larger values, so that the spine can be easily determined by local minima identification, as shown in Figure 20d. For cases of filaments crossing each other (forming objects in form of "X"), all branches longer than 31 pixels (minimum expected filament length) are reprocessed using the same technique, providing proper pruning as shown in Figure 21.

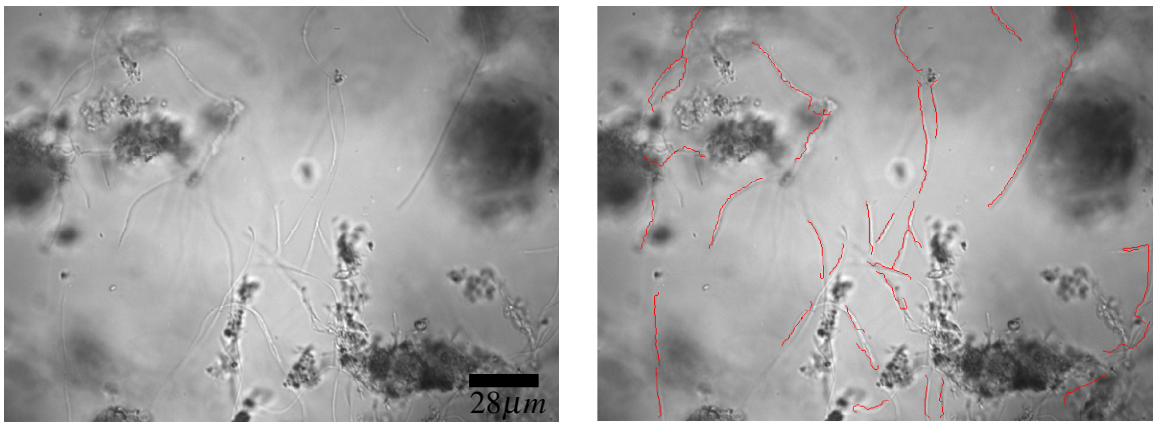
3.3.3.3 REDUCED RADIUS OF GYRATION (RRG)

At this stage, some objects originating from edges of larger agglomerates still need to be filtered out. A probabilistic approach based on the fact that true filaments are frequently more elongated and have less vertices than thin objects originating from

(a) Object skeleton. (b) Pruned skeleton.

Figure 21: Pruning technique applied to crossing filaments.

agglomerate edges is used. The Reduced Radius of Gyration (RRG) is calculated for the pruned objects to identify the true filaments (JENNÉ et al., 2001). Figure 23f provides a colormap illustrating the RRG values calculated for the detected pruned skeletons. Filaments are more elongated and therefore have larger RRG, so that via suitable thresholding they can be isolated as shown in Figure 22. As limitation, some short crossing filaments can present RRG values similar to non-filamentous objects.



(a) Original image. (b) Detected filaments.

Figure 22: Comparison between a) original ISM image and b) its version with detected filaments in red.

3.3.4 ESTIMATION OF FILAMENTS LENGTHS

Once the filaments are identified, their total extended length can be estimated. This information is obtained from the path lengths between endpoints already measured during pruning steps. Thereafter, their real length dimensions is estimated by applying a scale to these distance values. Given the pixel size of 6.45 x 6.45 μm^2, the image resize factor of a half and the $\times 40$ magnification, the scale factor is summarized as Equation 13.

$$length = distance \times \frac{2 \times 6.45}{40} [\mu m] \tag{13}$$

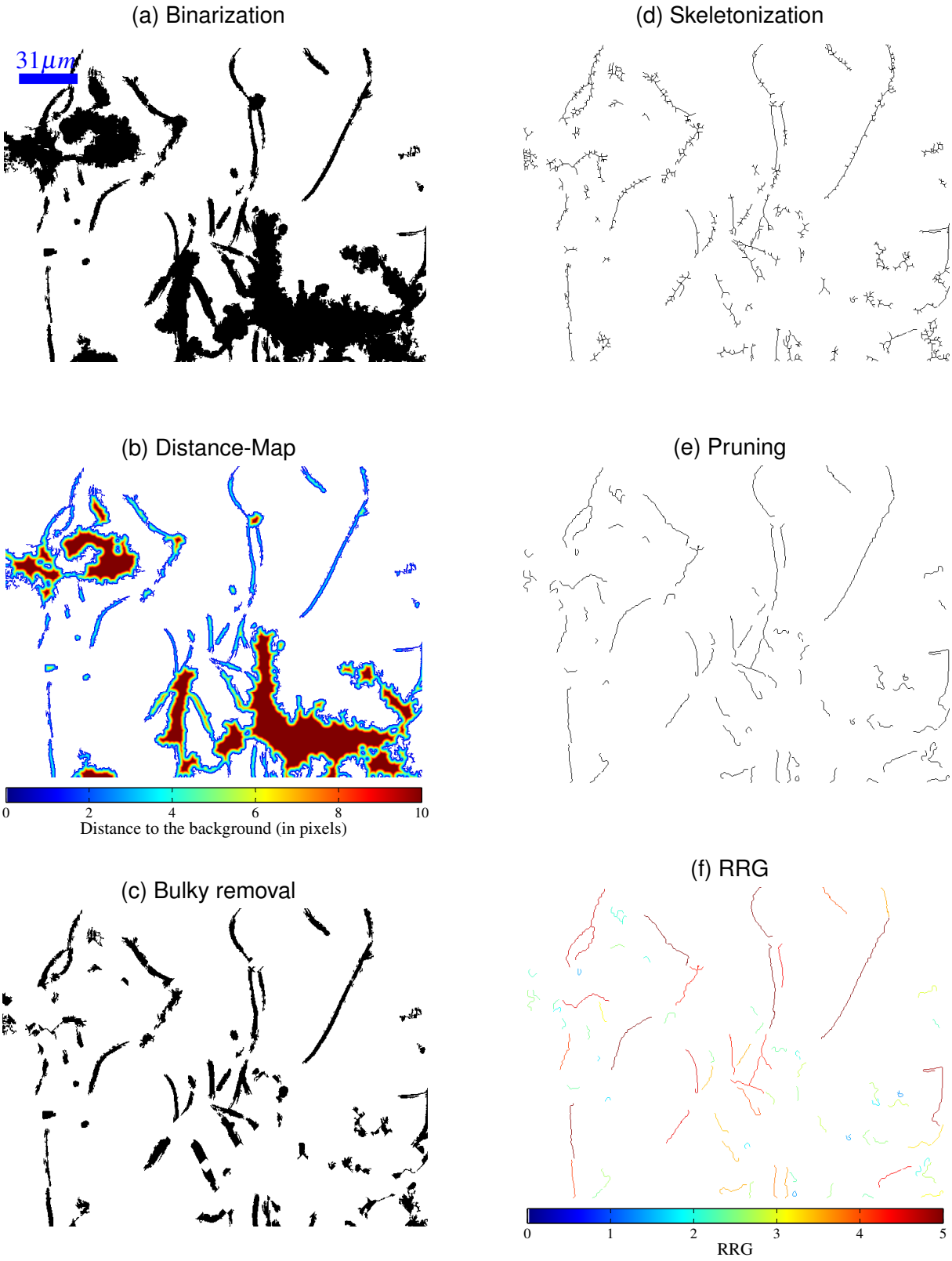


Figure 23: Images illustrating the sequence of operations for filament identification.

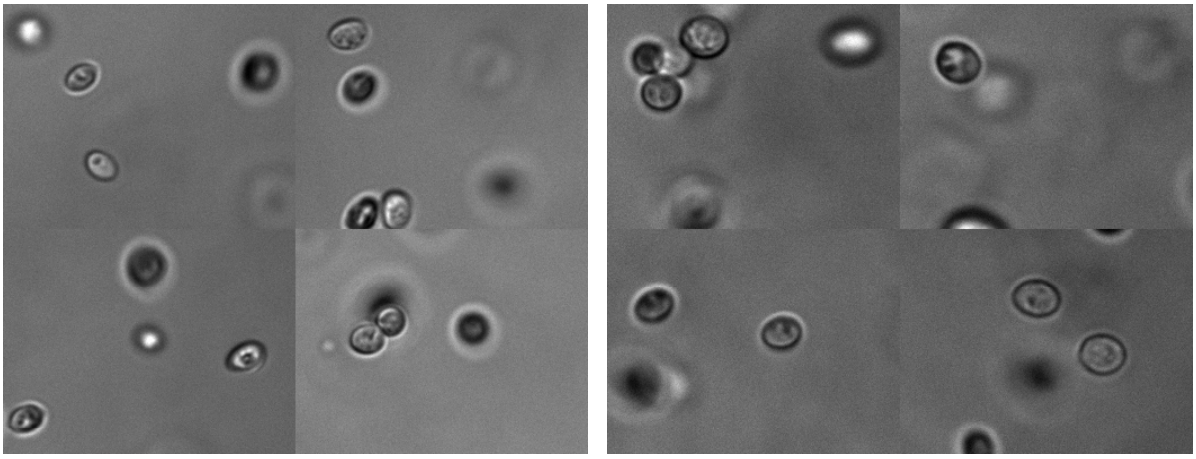
4 EXPERIMENTAL RESULTS AND DISCUSSION

Experiments were conducted to validate the two developments composing this project: i.e., new microscope design and algorithm for image processing. In section 4.1 are reported experiments performed to verify the image quality and the efficacy of the cleaning system. In section 4.2, comparisons between a reference (standard) method and the quantification of filaments by image processing are carried out to verify the accuracy of the new method.

4.1 IMAGE RESOLUTION

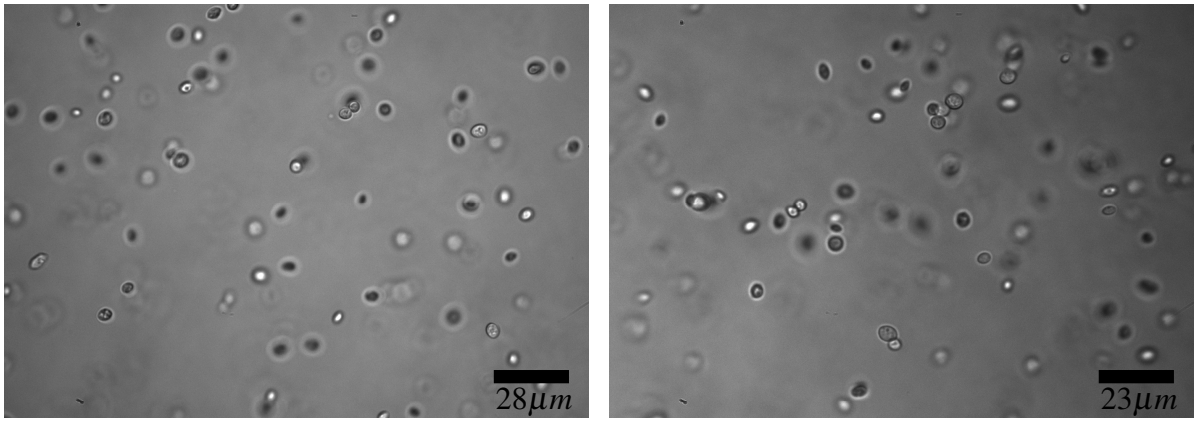
To evaluate the quality of the images acquired with the new ISM, the first experiment consisted in comparing images obtained with the new prototype (i.e., ISMv2) to the ones acquired with the ISMv1. A well-known application of *in situ* microscopy is the analysis of yeast cells, whose typical morphology can be observed in detail by ISM imaging (SUHR et al., 1995; CAMISARD et al., 2002). To evaluate the quality of the images provided by the new designed microscope, yeast cells were observed with both systems: ISMv1 and the 3D-printed prototype of the ISMv2. The quality of the obtained images were visually compared, as shown in Figures 24 and 25. The first figure is a composition of some sharp cells acquired with each system, while the second compares two single images acquired with each ISM.

One can observe that the new system ISMv2 presents a magnification larger than the ISMv1. This occurs because the new objective lens provides a $50\times$ magnification, against the $40\times$ provided by the previous objective lens. However, the resolution of the acquired images with the new system is very close to the resolution of the previous version, given that the yeast cells in focus appear sharply imaged, with internal details.



(a) Previous ISM (ISMv1) (b) New ISM (ISMv2)

Figure 24: Comparison – sharp cells with new ISM and its previous version. Magnification with the new system is larger, since the new objective provides a 50x magnification, against the 40x from the previous one. However, the resolution of the images obtained with both systems are quite similar, with some sharply imaged yeast cells (including internal details).

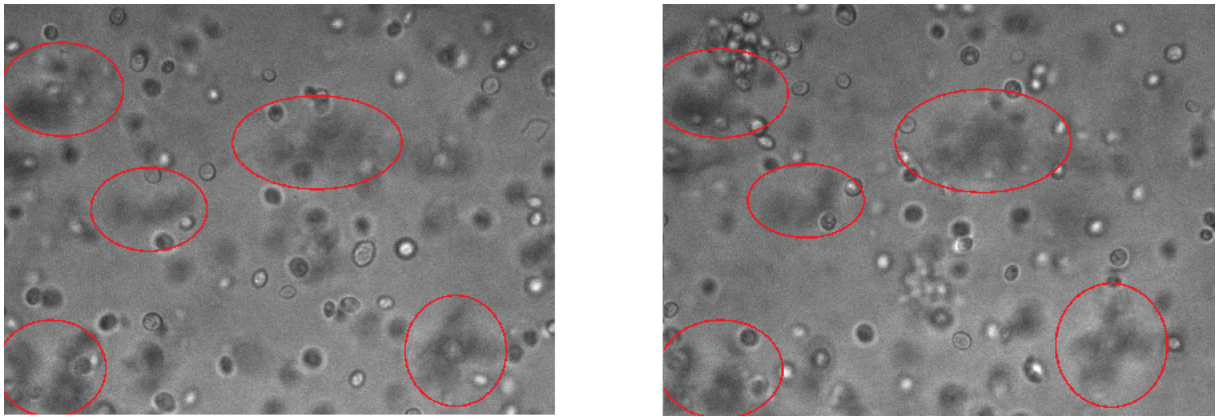


(a) Previous ISM (ISMv1) (b) New ISM (ISMv2)

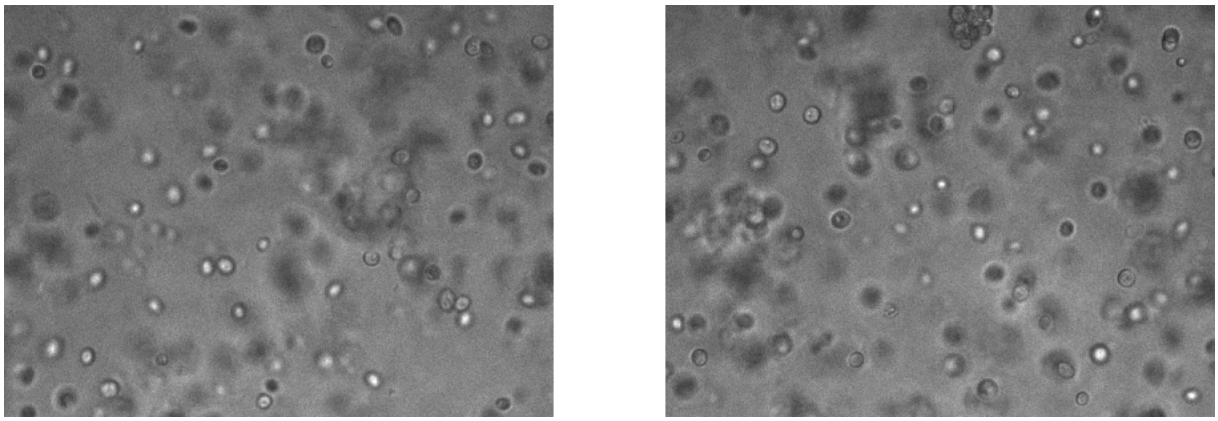
Figure 25: Comparison – images with a) new ISM and b) its previous version.

4.1.1 CLEANING SYSTEM

To test the magnetic cleaning system, the top part of the microscope was left overnight immersed in a dense suspension of yeast cells, so that some cells agglomerated and settled on the lens. An analogous experiment was also carried out using wastewater. Figure 26 and Figure 27 show sequences of images acquired during these experiments. In Figure 26, red circles highlight regions with settled yeast cells. In Figure 27, a long structure is immobile in the upper left corner before the cleaning action. These images prove the efficacy of the cleaning system: dirt initially present is removed and quality images can once again be acquired.



(a) Sequence of images before cleaning action



(b) Sequence of images after cleaning action

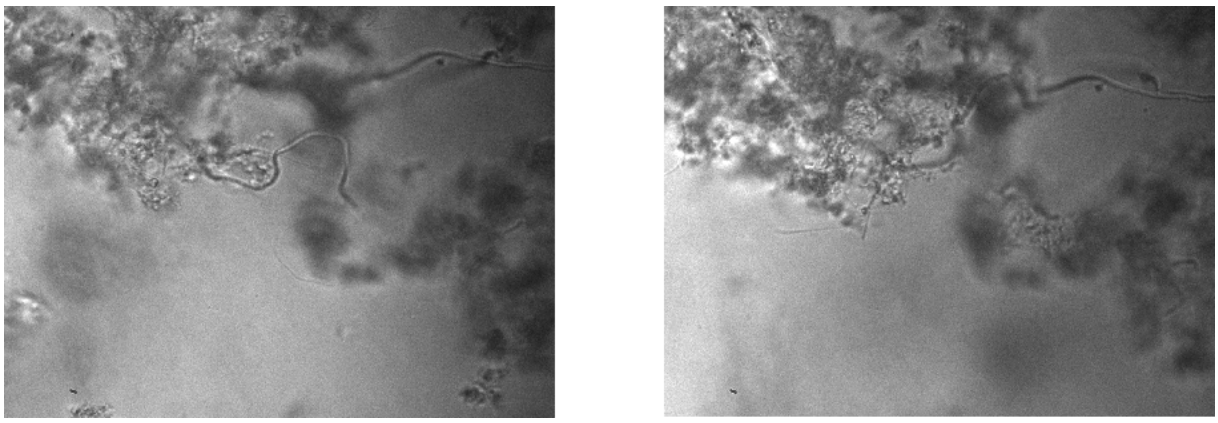
Figure 26: Sequence showing the cleaning system's action. a) Immobile cells due to adherence to the lens-surface are red circled. b) After cleaning action, all the cells previously settled were removed, what can be observed by comparing the circled regions.

4.2 QUANTIFICATION OF FILAMENTS

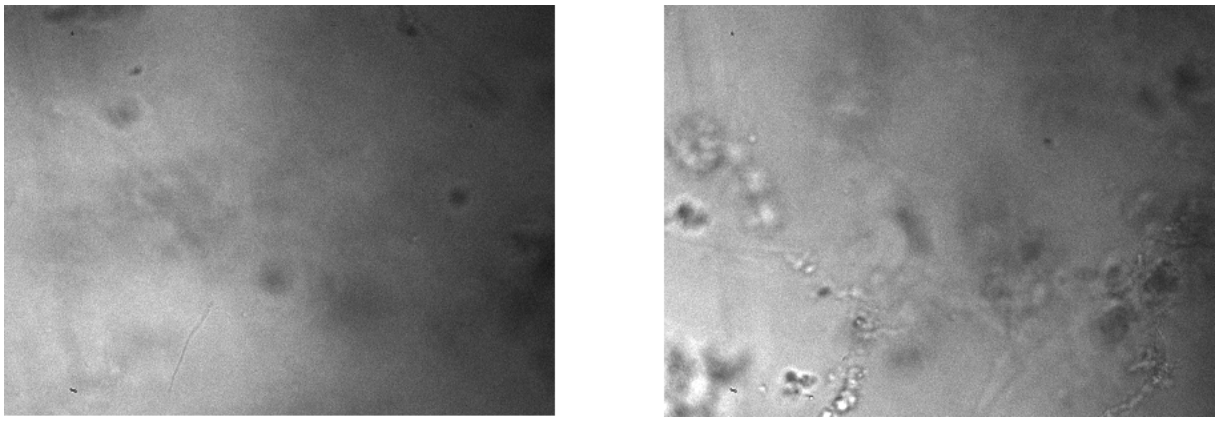
The applicability of the developed image processing algorithm for quantification of filamentous bacteria and estimation of their respective total extended filament length (TEFL) was verified through comparison to two different reference techniques.

4.2.1 EXPERIMENTAL SET-UP

Samples were weekly collected from an industrial activated sludge plant in Leverkusen, Germany (Currenta GmbH & Co OHG), from the second biological treatment step (cascade biology) and analyzed for a total period of twelve months. Results and further details of these experiments are explained in Dunkel et al.



(a) Before cleaning action



(b) After cleaning action

Figure 27: Sequence showing the cleaning system's action in wastewater. a) First, a long structure is immobile in the upper left corner, due to adherence to the lens-surface. b) After cleaning action, the structure is gone and proper imaging is possible.

(2015). For these initial experiments, the *in situ* microscope (ISM) was used off-line. Moreover, imaging was performed using the ISM previous version (ISMv1), since both developments – i.e., development of ISMv2 and conception of the image processing algorithm/analysis of samples – were performed in parallel.

Aiming at a set-up that properly replicates real environment conditions, samples were transported to laboratory and, without any prior conditioning as dilution or staining, directly introduced into a beaker. Inside this glass container, the ISM front-end sensor head was submerged and, for each sample, 500 images were acquired at a frequency of 3 images/second. Given the small virtual sample volume defined and since a magnet stir at 300 rpm kept the suspension agitated, all images are independent from each other, containing entirely new optical samples. As exemplified by Belini et al. (BELINI et al., 2013), at a flow speed of 0.1 m/s the samples are exchanged about 300 times per second.

4.2.2 HIT-RATIO OF FILAMENT DETECTION

Aiming at an independent evaluation of the image processing itself, without influences of bad imaging (i.e. low quality images due to configuration mistakes) or errors intrinsic to the actual reference methods, the hit-ratio of filament detection was computed using as reference twenty pre-marked images.

4.2.2.1 REFERENCE IMAGES

For evaluation of the image processing algorithm, real ISM images were selected from sets acquired for different activated sludge samples. The chosen reference images cover different scenarios, emphasizing difficult ones, such as: mix of filaments well and poorly focused, some connected to grainy structures; filaments crossing each other and grainy structures; short and long filaments; high concentration of flocs.

For establishment of the ground truth for filament detection, each reference image had its filaments marked by five researchers in the area of signal/image processing, who have a proper expertise to define which filaments are properly focused and should be detected by the algorithm. Filaments marked by three or more specialists were considered true filaments.

4.2.2.2 COMPUTING THE MAJORITY OF VOTES

The definition of reference filament markings by means of majority of votes was performed as follows:

1. first, all filament markings were dilated in order to compensate small displacements resulting from the filaments diameters and inaccuracies in manual marking;
2. an initial mask with all the pixels attribute to filaments was built by performing the *OR* logical operation with the images marked by the multiple specialists;
3. it was computed how many times each pixels in that mask was marked as filament. If more than half of the evaluators marked the referred pixel as filament, it was considered a true filament.

However, at this point the identified filaments are still dilated. For the further computation of filament detection accuracy, the original non-dilated forms of the reference filaments were necessary, since the amount of pixels composing each filament was used for this calculation. For their obtainment, their skeletonized version were compared to the skeletonized versions of all equivalent (i.e., in the same pixel coordinates) objects present in the images marked by each person. In this way was verified which originally marked filament was more similar (in amounts of coincident pixels) to the dilated one defined by majority of votes.

4.2.2.3 RECEIVER OPERATING CHARACTERISTICS (ROC) CURVES

The analysis of hit-ratio was performed using Receiver Operating Characteristics (ROC) curves (FAWCETT, 2006). They are two-dimensional graphs in which the True Positive ratio (TP) is plotted on the vertical axis, while the False Positive ratio (FP) is plotted on the horizontal one. Therefore, the upper left corner - i.e., the point $(0, 1)$ - corresponds to the perfect classification. Typically used for evaluation of hit rates in signal detection, ROC graphs have been also increasingly used in medical decision making, machine learning and data mining research (FAWCETT, 2006).

In the present study, true positives, false positives, true negatives and false negatives are defined as follows:

- Positives (P): pixels marked as filaments in the reference image;
- Negatives (N): pixels not marked as filaments in the reference image;
- True positives (TP): pixels marked as filaments in the reference image that were correctly classified by the algorithm as filaments;
- False negatives (FN): pixels marked as filaments in the reference image, but not detected by the algorithm;
- True negatives (TN): pixels neither identified by the algorithm as filaments nor marked in the reference image;
- False positives (FP): pixels identified by the algorithm as filaments, but not marked in the reference image;

As explained in Fawcett et al. (FAWCETT, 2006) and defined in Equation 14, the TP rate (TPR) is computed as the rate between the amount of elements identified

as true positives and the total number of positives in the ground truth – in our case, the total number of pixels marked as filaments in the reference image.

$$TPR = \frac{TP}{P} \tag{14}$$

The FP rate (FPR) is usually computed as the rate between FP and N. However, for the present application it would correspond to all other objects in the image (e.g. flocs and debris) and the whole background region. Thus, calculating the FPR as a ratio of the total number of negatives would result in a very low rate, providing misleading conclusions about the algorithm specificity. To proper evaluate the error introduced by the recognition of false filaments, we opted for calculating the FPR as a ratio of the total number of positives, expressed as:

$$FPR = \frac{FP}{P} \tag{15}$$

Figure 28 provides an example illustrating how TP, FN, TN and FP pixels are identified, for an object present in Figure 28a. The pixels marked as reference are shown in Figure 28b, while the ones detected by the algorithm are presented in Figure 28c.

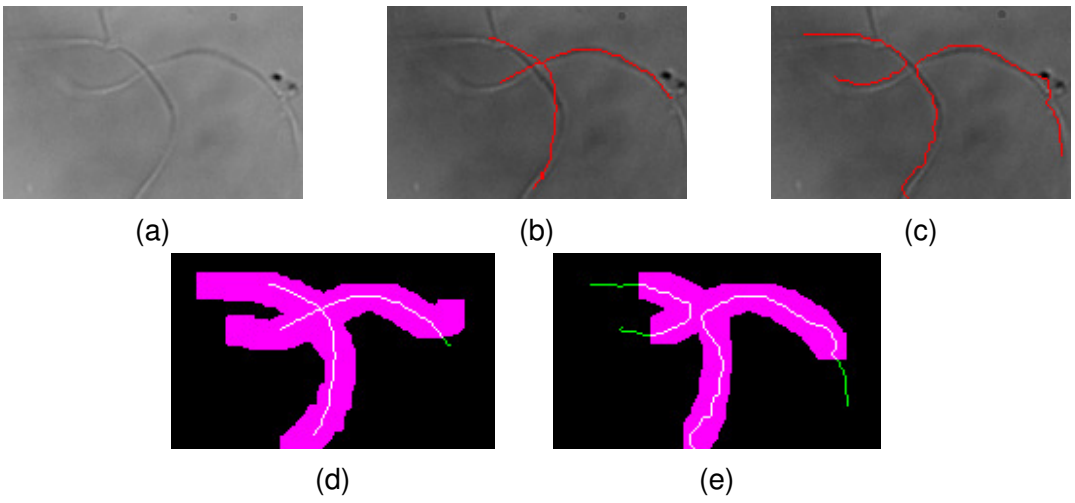


Figure 28: Steps for identification of pixels corresponding to TP, FN, TN and FP. a) Original object; b) Filaments marked by specialists; c) Filaments recognized by the proposed algorithm; d) Identification of TP (white) and FN (green); e) Identification of FP (green).

The recognition of true positives and false negatives pixels is illustrated in Figure 28d, where TP are in white and FN in green. It is performed through comparison between the reference markings and the dilated version of the algorithm detection, represented in purple. This dilation is necessary since small displacements can exist

between detected filaments and respective reference ones, due to their diameters and small imprecisions in manual marking. The recognition of false positives pixels, marked in green in Figure 28e, is performed through comparison between the algorithm detection and the dilated version of the reference markings.

4.2.2.4 OPTIMAL GAMMA CORRECTION

Using the methodology presented for computation of TP, FP, TN, FN parameters, pixelwise comparisons between reference filament markings and the ones identified by the algorithm were performed. For the gamma correction performed as preprocessing step, the value of 0.6 was defined based on the analysis of ROC-curves generated this way. Gamma values in the range 0.1-2.0 were evaluated, with the best results illustrated in Figure 29.

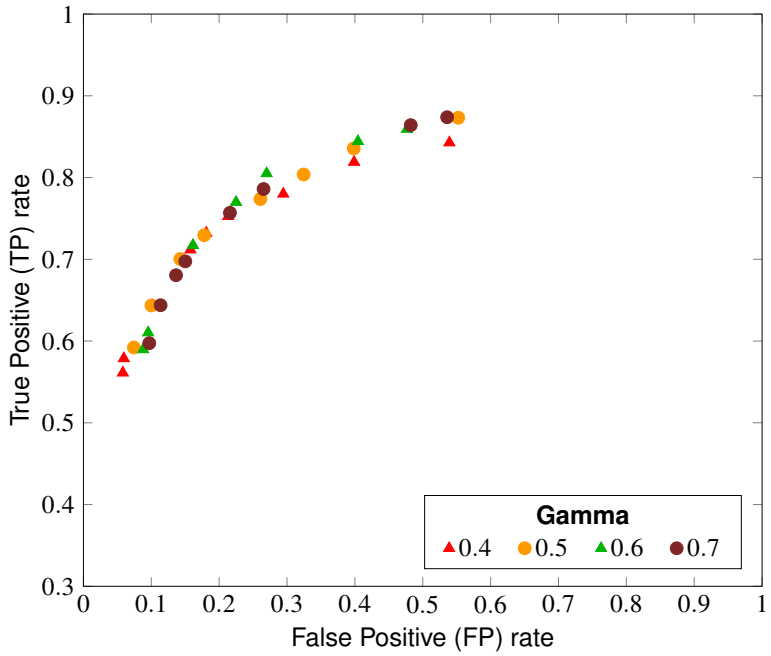


Figure 29: ROC curves with optimal parameters for gamma correction. Values within the range 0.1-2.0 were evaluated, with the best results here illustrated.

4.2.2.5 OPTIMAL VARIANCE, DISTANCE AND RRG THRESHOLDS

For optimization of variance, distance and RRG threshold values, ROC-curves were generated by computing multiple combinations. Each point of the ROC-curve shown in Figure 30 represents a different combination of variance, distance and RRG

thresholds. This chart illustrates how the variance threshold value acts as the coarse adjustment for the algorithm's hit-ratio. Points corresponding to combinations with a common variance threshold value form well-defined groups. From this first analysis, it is observed that the best results are obtained with variance threshold values between 8×10^{-4} and 12×10^{-4}.

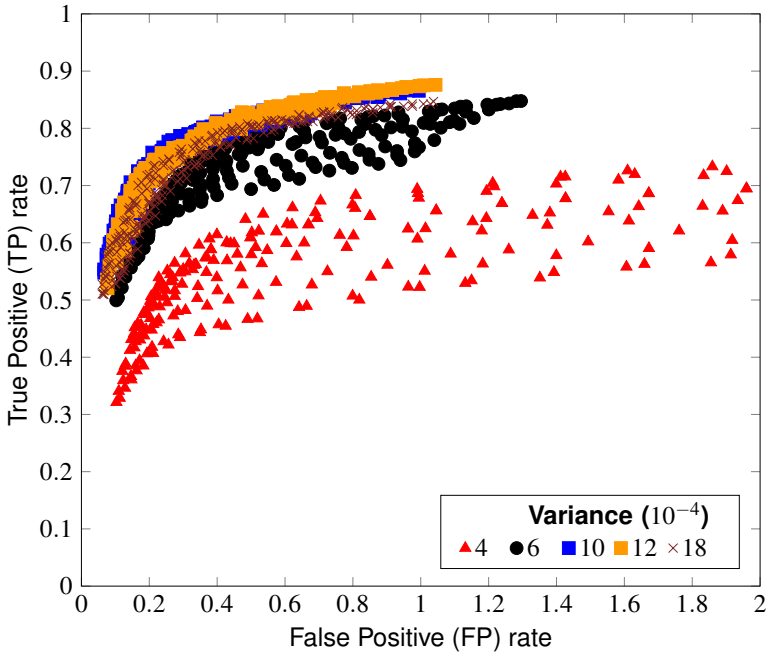


Figure 30: ROC curve with different combinations of variance, distance and RRG threshold values. Markers/colors are assigned according to variance threshold. Combinations with a common variance threshold form well-defined groups, so that the variance corresponds to the algorithm's coarse adjustment.

Source: Adapted from (DIAS et al., 2016).

Meanwhile, the values of distance and RRG thresholds correspond to finer adjustments inside each of these groups. As illustrated in Figure 31 for one variance group (6×10^{-4}), combinations sharing common distance threshold values are distributed in form of well-defined curves inside each variance group. Higher distance thresholds provide higher TPR at the expense of a respective FPR, since it implies that only larger objects will be classified as flocs and thereafter filtered out. Finally, Figure 32 shows that the values of RRG threshold correspond to a latter fine adjustment which also defines the system's sensibility at the expense of a respective specificity.

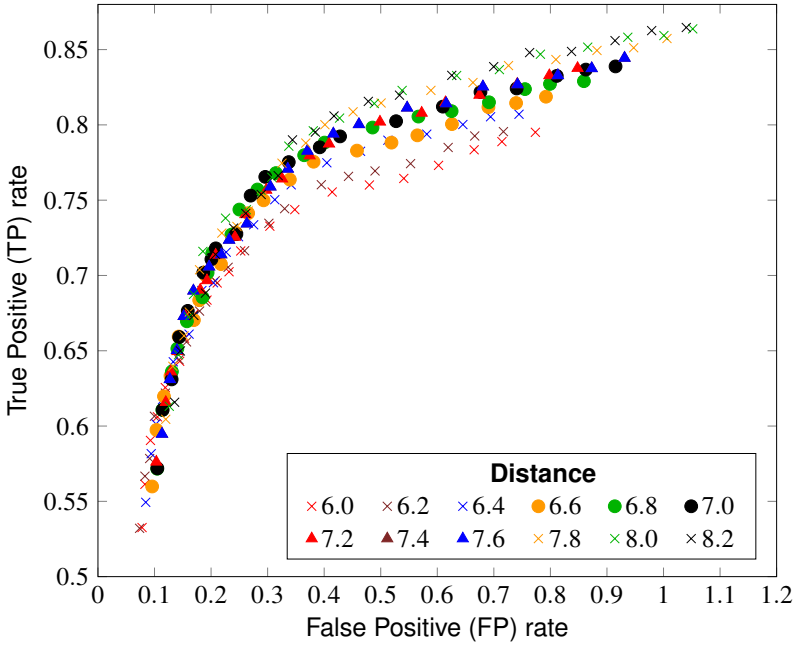


Figure 31: ROC curve for variance threshold 6×10^{-4}. Combinations with common variance and distance threshold values form well-defined curves, being the distance threshold a finer adjustment.

Source: Adapted from (DIAS et al., 2016).

The ROC-curve in Figure 33 contains only points of optimal combinations. Points composing the curve's convex-hull are marked in black, with their respective thresholds, TPR (equivalent to sensitivity), FPR and specificity ratios shown in Table 5. In addition, the same chart also illustrates the hit-ratio obtained for two different groups: in blue, the results obtained for a group composed by the ten images for which the TPR values obtained were higher than the median; in red, values obtained for the group composed by images that yielded TPR lower than the median.

Results revealed that, for these reference images, the automated analysis in average correctly detected up to 72% of the pixels marked by specialists as filaments, with a false-positive ratio equivalent to 14% of this value (point 6 in Figure 33). This corresponds to a specificity of about 86%. Moreover, as illustrated by the blue and red curves, the true positive ratio oscillates between 64% and 81% for this optimal combination of thresholds, with a false positive ratio between 11% and 17%.

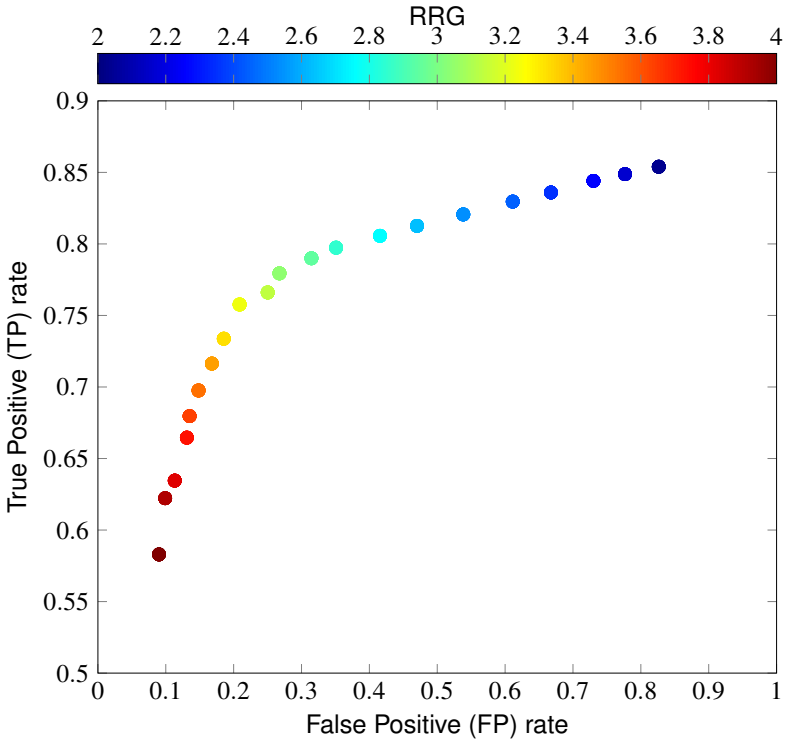


Figure 32: ROC curve for fixed variance and distance threshold values (10×10^{-4} and 7.0, respectively), with a colormap illustrating different RRG threshold values. The RRG acts as the algorithm's final adjustment.

Source: Adapted from (DIAS et al., 2016).

Table 5: Best combinations of thresholds composing the ROC curve

Point	Variance (10^{-4})	Distance	RRG	TPR [%]	FPR [%]	Specificity
1	12.0	7.0	2.6	82.82	47.31	0.5269
2	9.5	8.0	3.1	80.42	31.54	0.6846
3	9.5	8.0	3.2	78.87	27.33	0.7267
4	10.0	7.0	3.2	75.77	20.88	0.7912
5	10.5	6.4	3.2	72.31	14.38	0.8562
6	10.5	6.3	3.2	70.74	13.74	0.8626
7	11.5	6.6	3.4	68.39	12.29	0.8771
8	10.5	6.3	3.5	65.33	10.19	0.8981
9	9.5	6.4	3.7	62.63	8.56	0.9144
10	9.5	6.4	3.9	59.34	7.28	0.9272
11	9.5	6.2	4.0	55.80	6.07	0.9393

Source: Extracted from (DIAS et al., 2016).

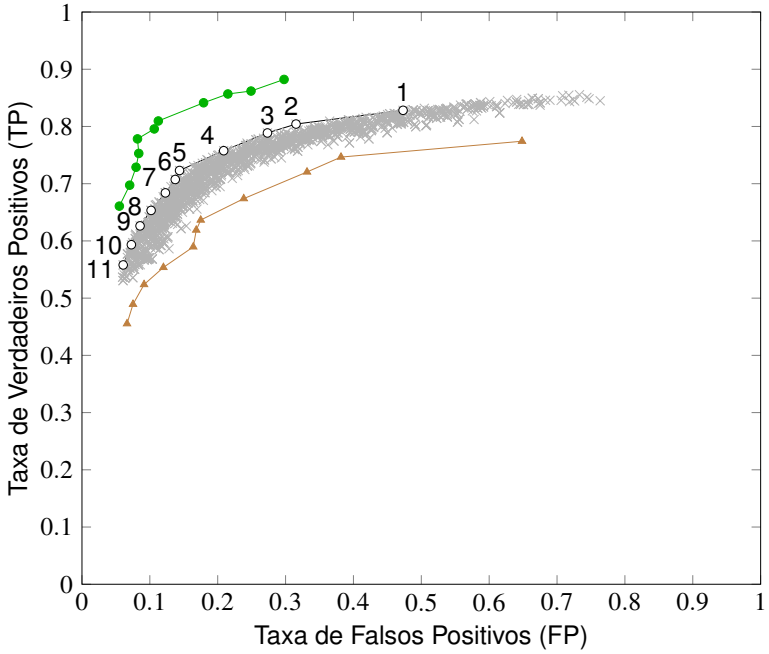


Figure 33: ROC curve of algorithm's hit-ratio with the best combinations of variance, distance and RRG threshold values. Optimal points along the curve's convex-hull are illustrated in black, with numbers assigned for matching with Table 6. In blue, results obtained for the ten images which yielded TPR higher than the median; in red, results obtained for the images that yielded TPR lower than the median.

Source: Adapted from (DIAS et al., 2016).

4.2.3 TOTAL EXTENDED FILAMENT LENGTH (TEFL)

One of the standard methods for quantification of filaments is the total extended filament length (TEFL), calculated according to Sezgin et al. (1978). For comparison with the reference TEFL method, 500 images from each wastewater sample were acquired and processed, so that an average ISM-online TEFL (or ISM-oTEFL) per image is computed by the image processing algorithm here proposed. As described by Dunkel et al. (2015) and illustrated in Figure 34, a linear relationship between the two methods is indicated by a Pearson correlation coefficient of $r = 0.87$. That means the proposed system is capable of detecting filamentous bacteria with good correlation to the reference TEFL method, specially in terms of trends over time, which is the critical information for decision taking against overgrowth of filamentous bacteria in activated sludge. Moreover, the suitability of the developed system for monitoring sludge settling properties was also supported by a Granger causality test, which yielded a p-value of 0.0026 for the relationship of ISM-oTEFL and the Diluted Sludge Volume Index (DSVI) (STÖBBE, 1964), a parameter that expresses sludge settling characteristics.

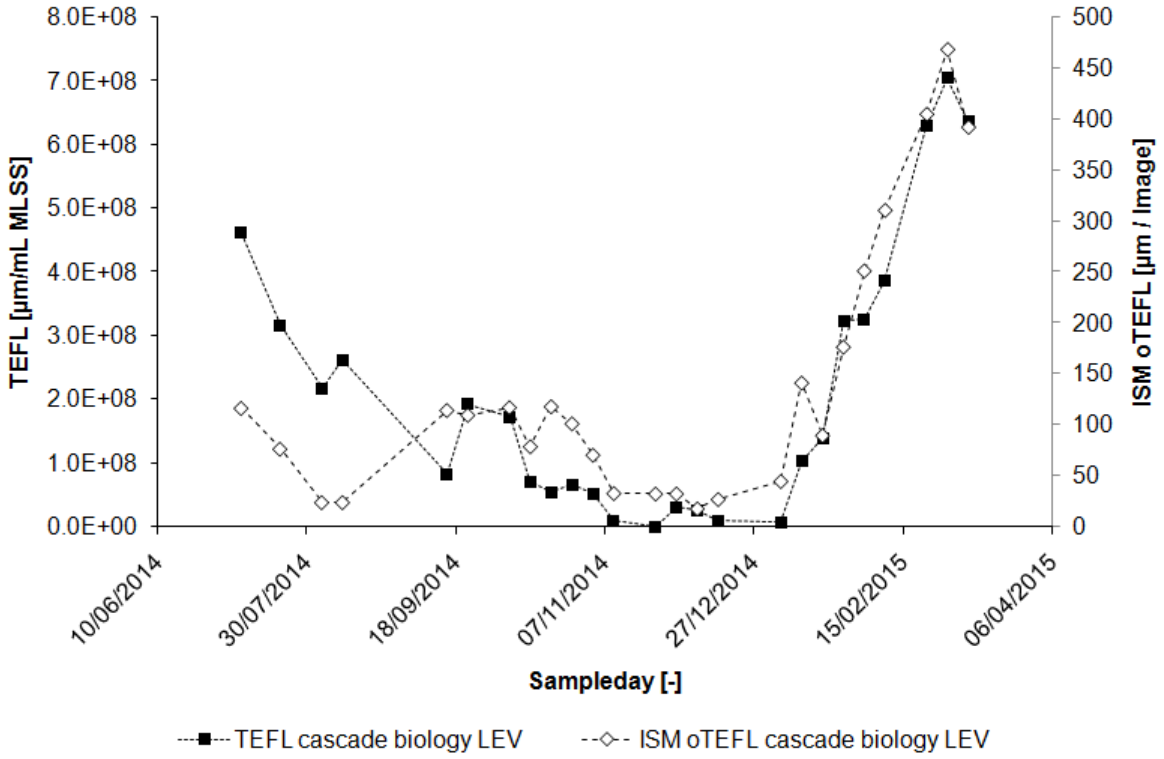


Figure 34: Comparison between TEFL and ISM-oTEFL. A good correlation ($r = 0.87$) is observed, specially in terms of trends over time.

Source: Extracted from (DUNKEL et al., 2015).

5 CONCLUSION

The current standard methods for quantification of filamentous bacteria are based on discontinuous microscopy and visual analysis by human operators, so that they are labor-intensive and information can be lost during sampling, transportation and preparation of samples. Moreover, a limited number of images is analysed, temporal changes are difficult to be examined and the techniques are subjective and susceptible to human errors.

Aiming a more reliable, consistent method, this project proposes the application of *in situ* microscopy for the monitoring of filamentous bacteria in wastewater treatment plants. A proper solution for the proposed problem required two distinct tasks: adjustments in the microscope design and the development of a suitable image processing algorithm, since no staining or phase-contrast techniques are employed.

A new version of the ISM was modeled in AutoCAD, resulting in a simpler system, with easier and safer operation. This was possible through the use of a new objective lens that allowed the design of a single tube microscope. Essentially composed by three commercial lenses, such objective lens designed by Professor Hajo Suhr in collaboration with the University of Stuttgart can be immersed directly in the suspension. Nevertheless, the use of the ISM in a wastewater environment required the development of a cleaning system, whose operation is based in magnetic principles to avoid extra physical connections traversing the ISM. In this way, impermeabiliy is ensured and contamination of the environment avoided.

To verify the suitability of the new ISM, a prototype was built using a 3D-Printer. For quality evaluation of the images obtained with this new system, yeast cells were imaged with both the previously developed ISM and the newly prototyped one. Visual inspection confirmed that the new system provides images with similar resolution to the ones obtained with the former microscope. For example, even yeast cells internal structures in the order of μm are sharply imaged. The cleaning system developed could also successfully remove particles deliberately settled above lenses, reestablishing the acquisition of quality images.

Experiments have shown that the developed image processing algorithm can detect around 72% of the amount of pixels marked by specialists as filaments in reference images, with a false positive rate of 14% (i.e., specificity of about 86%). Moreover, it is capable of detecting filamentous bacteria with good correlation to the reference TEFL method, specially in terms of trends over time, which is the critical information for plant operators actions against overgrowth of filamentous bacteria in activated sludge. With an average execution time of 0.7 second per image, the algorithm is therefore suitable for being optimally mapped into a computational architecture to provide real-time monitoring.

In summary, it is possible to affirm that the objectives defined for this master's research project were achieved. A new simplified ISM, with cleaning system and capable of imaging with compatible resolution was designed and the proposed algorithm for image processing provided quantification of filaments well correlated to the ones provided by standard techniques. In addition, this project have led to a conference poster (DUNKEL et al., 2014) and an article as second author in the Journal of Water Science and Technology (DUNKEL et al., 2015), with another one focused on the developed image processing algorithm waiting for editor review after its recent submission.

5.1 FUTURE WORK

Future work includes evaluation under real conditions by attaching an *in situ* microscope into a pilot wastewater treatment plant, avoiding possible errors caused by sampling, storage and transporting. In addition, real-time analysis could be carried out, facilitating correlations to the other standard biological indicators. The production of metallic prototypes is also important to validate the reproducibility and ease of assembly of the new system, specially the new objective lens and cleaning system composing it. Furthermore, the ISM can be used in the future as an automatic controlling tool for specific counter measures against bulking and foaming.

For the ISM used in this work, only microbial aggregates smaller than 0.3 mm were able to pass through the gap between the fiber-ending and the quartz glass window separating objective and suspension. Microscopic analyses have shown that the samples evaluated up to now were composed by mainly stable and compact sludge flocs, with maximal size of 241.6 μm (DUNKEL et al., 2016). If future experiments should prove that a bigger gap is necessary, both microscope and image processing

algorithm can be adapted, since it would be a matter of brightness/gain and thresholds adjustments.

The proposed algorithm has also room for improvements, mainly for cases such as short crossing filaments which are eliminated by the RRG filtering. Besides, closely spaced filaments are frequently binarized as a single, massive object that is further removed by the distance filter.

A stand-alone software is desirable, which involves converting the algorithm written in MATLAB into OpenCV or another similar open source tool for image processing. Furthermore, an user interface also needs to be developed for visualization and processing of data, in addition to possible adjustments in parameters settings.

REFERENCES

AMARAL, A.; FERREIRA, E. Activated sludge monitoring of a wastewater treatment plant using image analysis and partial least squares regression. **Analytica Chimica Acta**, vol. 544, n. 1-2, p. 246–253, 2005. ISSN 00032670.

BELINI, V. L.; WIEDEMANN, P.; SUHR, H. In situ microscopy: A perspective for industrial bioethanol production monitoring. **Microbiological Methods**, vol. 93, p. 224–232, 2013.

BISCAIA, M. F. **Embedded System and User Interface for an In Situ Microscope**. Dissertation (Master) — Federal University of Technology - Paraná and Hochschule Mannheim, 2011.

BITTNER, C.; WEHNERT, G.; SCHEPER, T. In situ microscopy for on-line determination of biomass. **Biotechnology and Bioengineering**, Wiley, Inc., vol. 60, n. 1, p. 24–35, 1998.

BITTON, G. **Wastewater Microbiology**. New York, NY, USA: Wiley, 2005.

CAMISARD, V.; BRIENNE, J.-P.; BAUSSART, H.; HAMMANN, J.; SUHR, H. Inline characterization of cell concentration and cell volume in agitated bioreactors using in situ microscopy: Application to volume variation induced by osmotic stress. **Biotechnology and Bioengineering**, Wiley Subscription Services, Inc., vol. 78, n. 1, p. 73–80, 2002.

CENENS, C.; BEURDEN, K. V.; JENNÉ, R.; IMPE, J. V. On the development of a novel image analysis technique to distinguish between flocs and filaments in activated sludge images. **Water Science and Technology**, vol. 46(1-2), p. 381–387, 2002.

COSTA, J. C.; MESQUITA, D. P.; AMARAL, A. L.; ALVES, M. M.; FERREIRA, E. C. Quantitative image analysis for the characterization of microbial aggregates in biological wastewater treatment: a review. **Environmental Science and Pollution Research**, vol. 20, n. 9, p. 5887–5912, 2013. ISSN 0944-1344.

Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V - DWA. **Leistungsvergleich kommunaler Kläranlagen 2011**. Hennef (Germany), 2011.

DIAS, P. A.; DUNKEL, T.; FAJADO, D. A. S.; GALLEGOS, E. d. L.; DENECKE, M.; WIEDEMANN, P.; SCHNEIDER, F. K.; SUHR, H. Image processing for identification and quantification of filamentous bacteria in in situ acquired images. **BioMedical Engineering OnLine**, vol. 15, n. 1, p. 1–19, 2016. ISSN 1475-925X.

DUNKEL, T.; DIAS, P. A.; GALLEGOS, E. L. de L.; TACKE, V.; SCHIELKE, A.; HESSE, T.; FAJADO, D. A. S.; SUHR, H.; WIEDEMANN, P.; DENECKE, M. In situ microscopy as a tool for the monitoring of filamentous bacteria: a case study in an industrial activated sludge system dominated by *M. parvicella*. **Water Science and Technology**, IWA Publishing, 2015. ISSN 0273-1223.

DUNKEL, T.; GALLEGOS, E. L. de L.; SCHÖNSEE, C. D.; HESSE, T.; JOCHMANN, M.; WINGENDER, J.; DENECKE, M. Evaluating the influence of wastewater composition on the growth of *Microthrix parvicella* by GCxGC/qMS and real-time PCR. **Water Research**, vol. 88, p. 510 – 523, 2016. ISSN 0043-1354.

DUNKEL, T.; LEON, E. D.; DIAS, P.; SUHR, H.; DENECKE, M. In situ microscopy - early warning system for the growth of filamentous bacteria. In: IWA. **Poster at: IWA Conference Activated Sludge - 100 Years and Counting**. Essen (Germany), 2014.

Economic and Social Commision for Western Asia. **Wastewater treatment technologies: a general review**. New York; NY, USA, 2003.

EIKELBOOM, D. **Process Control of Activated Sludge Plants by Microscopic Investigation**. London, UK: IWA Publishing, 2000.

FAWCETT, T. An introduction to ROC Analysis. **Pattern Recognition Letters**, Elsevier Science Inc., New York, NY, USA, vol. 27, n. 8, p. 861–874, jun. 2006. ISSN 0167-8655.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 3rd. ed. Upper Saddle River, NJ, USA: Prentice Hall, 2007. Hardcover.

GUEZ, J.; CASSAR, J.; WARTELLE, F.; DHULSTER, P.; SUHR, H. Real time in situ microscopy for animal cell-concentration monitoring during high density culture in bioreactor. **Biotechnology**, vol. 111, p. 335–343, 2004.

GUEZ, J.; CASSAR, J.; WARTELLE, F.; DHULSTER, P.; SUHR, H. The viability of animal cell cultures in bioreactors: Can it be estimated online by using in situ microscopy? **Process Biochemistry**, 2009.

Instituto Brasileiro de Geografia e Estatística. **Pesquisa Nacional de Saneamento Básico**. Rio de Janeiro, Rio de Janeiro, Brazil, 2008.

JENKINS, D.; RICHARD, M.; DAIGGER, G.; COMMISSION, S. A. W. R. **Manual on the Causes and Control of Activated Sludge Bulking and Foaming**. London, UK: Water Research Commission, 1984.

JENKINS, D.; RICHARD, M. G.; DAIGGER, G. T. **Manual on the causes and control of activated sludge bulking, foaming, and other solids separation problems**. 3rd. ed. Hoboken, NJ, USA: CRC Press, 2003.

JENNÉ, R.; BANADDA, E. N.; SMETS, I.; DEURINCK, J.; IMPE, J. V. Detection of filamentous bulking problems: developing an image analysis system for sludge composition monitoring. **Microscopy and microanalysis**, vol. 13, p. 36–41, 2007. ISSN 1431-9276.

JENNÉ, R.; CENENS, C.; IMPE J. F, . V. Towards on-line quantification of flocs and filaments by means of image analysis for optimization and control of activated sludge plants. **Forum for Applied Biotechnology (FAB)**, vol. 66 (3B), p. 63–70, 2001.

KHAN, M. B.; LEE, X. Y.; NISAR, H.; NG, C. A.; YEAP, K. H.; MALIK, A. S. Digital image processing and analysis for activated sludge wastewater treatment. **Signal and Image Analysis for Biomedical and Life Sciences**, Springer International Publishing, Cham, p. 227–248, 2015.

KOIVURANTA, E.; KESKITALO, J.; HAAPALA, A.; STOOR, T.; SARÉN, M.; NIINIMÄKI, J. Optical monitoring of activated sludge flocs in bulking and non-bulking conditions. **Environmental Technology**, vol. 34, n. 5, p. 679–686, 2013. ISSN 0959-3330.

KUNST, S.; HELMER, C.; KNOOP, S. **Betriebsprobleme auf Klaeranlagen durch Blaehschlamm, Schwimmschlamm, Schaum - Handbuch zur Identifizierung und Bekaempfung faediger Bakterien**. Berlin, Germany: Springer, 2000.

LOPEZ, C.; PONS, M. N.; MORGENROTH, E. Evaluation of microscopic techniques as a basis for the quantitative image analysis of activated sludge. **Water Research**, vol. 39, n. 2-3, p. 456–468, 2005. ISSN 00431354.

MARQUES, O. **Pratical Image and Video Processing Using MATLAB**. Hoboken, NJ, USA: John Wiley & Sons, 2011.

MESQUITA, D.; AMARAL, A.; FERREIRA, E. Identifying different types of bulking in an activated sludge system through quantitative image analysis. **Chemosphere**, vol. 85, p. 643–652, 2011.

MESQUITA, D. P.; DIAS, O.; AMARAL, A. L.; FERREIRA, E. C. A Comparison between bright field and phase-contrast image analysis techniques in activated sludge morphological characterization. **Microscopy and microanalysis**, vol. 16, n. 2, p. 166–74, 2010. ISSN 1435-8115.

MOTTA, M. da; PONS, M.-N.; ROCHE, N.; VIVIER, H. Characterisation of activated sludge by automated image analysis. **Biochemical Engineering Journal**, vol. 9, n. 3, p. 165 – 173, 2001. ISSN 1369-703X.

PONS, M. N.; VIVIER, H. Biomass quantification by image analysis. **Advances in biochemical engineering/biotechnology**, vol. 66, p. 133–184, 2000. ISSN 0724-6145 (Print).

RUSS, J. C. **Image Processing Handbook**. 4th. ed. Boca Raton, FL, USA: CRC Press, Inc., 2002.

SCHOLZ, T.; JAHNE, B.; SUHR, H.; WEHNERT, G.; GEISSLER, P.; SCHNEIDER, K. A new depth from focus technique for in situ determination of cell concentration in bioreactors. In: DAGM. **Proc. 16. DAGM-Symposium Mustererkennung**. Vienna, Austria, 1994. p. 145–150.

SEVIOUR, R.; NIELSEN, P. **Microbial Ecology of Activated Sludge**. London, UK: IWA Publishing, 2010.

SEZGIN, M.; JENKINS, D.; PARKER, D. S. A unified theory of filamentous activated sludge bulking. **Water Pollution Control Federation**, vol. 50, No. 2, p. 362–381, 1978.

SOILLE, P. **Morphological Image Analysis: Principles and Applications**. 2. ed. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2003. ISBN 3540429883.

STÖBBE, G. Über das Verhalten von Belebtschlamm in Aufsteigender Wasserbewegung. **Veröffentlichungen der Technischen Hochschule Hannover**, vol. 18, 1964.

Hajo Suhr, Peter Speil, Gerd Wehnert and Winfried Storhas. **Bioreaktor mit In Situ Mikroskopsonde und Messverfahren**. 1991. DE-4032002.

SUHR, H.; WEHNERT, G.; SCHNEIDER, K.; BITTNER, C.; SCHOLZ, T.; GEISSLER, P.; JÄHNE, B.; SCHEPER, T. In situ microscopy for on-line characterization of cell-populations in bioreactors, including cell-concentration measurements by depth from focus. **Biotechnology and Bioengineering**, Wiley Inc., vol. 47, n. 1, p. 106–116, 1995.

TCHOBANOGLOUS, G.; BURTON, F.; STENSEL, H. **Wastewater Engineering: Treatment and Reuse**. Boston, USA: McGraw-Hill Education, 2003.

THE MATHWORKS, INC. **MATLAB and Statistics Toolbox Release 2013a**. Natick, Massachusetts, United States, 2013. Available at <http://www.mathworks.com/help/matlab/>.

University of the Basque Country. **In silico simulation of molecular biology experiments**. Leioa, Bizkaia, Spain, 2015. Available at <http://insilico.ehu.eus/counting_chamber/thoma.php/>.

WIEDEMANN, P.; GUEZ, J.; WIEGEMANN, H. B.; EGBER, F.; QUINTANA, J. C.; ASANZA-MALDONADO, D.; FILIPAKI, M.; WILKESMAN, J.; SCHWIEBERT, C.; CASSAR, J. P.; DHULSTER, P.; SUHR, H. In situ microscopic cytometry enables noninvasive viability assessment of animal cells by measuring entropy states. **Biotechnology and Bioengineering**, 2011.

WIEDEMANN, P.; WORF, M.; WIEGEMANN, H.; EGNER, F.; SCHWIEBERT, C.; WILKESMAN, J.; GUEZ, J.; QUINTANA, J.; ASSANZA, D.; SUHR, H. On-line and real time cell counting and viability determination for animal cell process monitoring by in situ microscopy. **BMC Proceedings**, vol. 5, n. 8, p. P77, 2011.

WWAP - United Nations World Water Assessment Programme. **The United Nations World Water Development Report 2014: Water and Energy**. Paris, 2014.