ELSEVIER

Journal of Microbiological Methods 47 (2001) 169–180

Journal of Microbiological Methods

www.elsevier.com/locate/jmicmeth

Objective threshold selection procedure (OTS) for segmentation of scanning laser confocal microscope images

J.B. Xavier [a], A. Schnell [b], S. Wuertz [b], R. Palmer [c], D.C. White [d,e], J.S. Almeida [a,f,*]

[a] *ITQB / UNL, R Qta Grande 6, 2780 Oeiras, Portugal*

[b] *Institute of Water Quality Control and Waste Management, Technical University of Munich, D-85748 Garching, Germany*

[c] *National Institute of Dental Craniofacial Research-National Institutes of Health, Oral Infection and Immunity Branch, Bldg. 30, Room 308, 30 Convent Drive, Bethesda, MD 20892, USA*

[d] *Center for Environmental Biotechnology, The University of Tennessee, 10515 Research Drive, Suite 300, Knoxville, TN 37932, USA*

[e] *Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*

[f] *Department of Biometry and Epidemiology, Medical University of South Carolina, 135 Rutledge Avenue, PO Box 250551, Charleston, SC 29425, USA*

Received 27 March 2001; received in revised form 12 June 2001; accepted 18 June 2001

Abstract

The determination of volumes and interface areas from confocal laser scanning microscopy (CLSM) images requires the identification of component objects by segmentation. An automated method for the determination of segmentation thresholds for CLSM imaging of biofilms was developed. The procedure, named objective threshold selection (OTS), is a three-dimensional development of the approach introduced by the popular robust automatic threshold selection (RATS) method. OTS is based on the statistical properties of local gray-values and gradients in the image. By characterizing the dependence between a volumetric feature and the intensity threshold used for image segmentation, the former can be determined with an arbitrary confidence level, with no need for user intervention. The identification of an objective segmentation procedure renders the possibility for the full automation of volume and interfacial area measurement. Images from two distinct biofilm systems, acquired using different experimental techniques and instrumental setups were segmented by OTS to determine biofilm volume and interfacial area. The reliability of measurements for each case was analyzed to identify optimal procedure for image acquisition. The automated OTS method was shown to reproduce values obtained manually by an experienced operator.

Keywords: Objective threshold selection; Confocal laser scanning microscopy; Robust automatic threshold selection

1. Introduction

Confocal laser scanning microscopy (CLSM) is the method of choice to capture structure of live biofilms due to its non-invasive and non-destructive character. Series of optical cross-sections, collected at different depths in order to scan a given volume of interest, provide extensive three-dimensional structural data. Time-course analysis of biofilm morphogenesis is also possible using CLSM, since it allows for non-destructive and repetitive visualization of live biological structures in their natural hydrated state (Lawerence et al., 1991). A number of struc-

* Corresponding author. Department of Biometry and Epidemiology, Medical University of South Carolina, 135 Rutledge Avenue, PO Box 250551, Charleston, SC 29425, USA.

0167-7012/01/$ - see front matter
PII: S0167-7012(01)00298-6

ture/function studies of biofilm formation have been published over the past decade that use qualitative analysis of CLSM images (Lawerence et al., 1991; DeBeer et al., 1994; Lawerence et al., 1994; Caldeira et al., 1999; Sternberg et al., 1999; Wood et al., 1998; Jayaraman et al., 1998). Quantitative image analysis has also been used, but has so far been either limited to two dimensions (Yang et al., 1999; Hermanowicz et al., 1995) or required operator-defined calibration (Jayaraman et al., 1998; Hermanowicz et al., 1995; Kuehn et al., 1998; Stoodley et al., 1998; Christensen et al., 1998; Moller et al., 1998). The full potential of CLSM for quantitative study of biofilms has not been fully automated due to two main factors: the large computational requirements for the processing of three-dimensional data and the unavailability of adequate mathematical solutions and the corresponding computational tools.

Biofilm structural parameters such as distribution of density and surface shape are well recognized in being of key importance for the stability and performance of biofilm reactors (Picioreanu et al., 1997). Nevertheless, biofilm processes are typically modeled under the assumption that these are uniform cell aggregates that homogeneously cover a solid surface (Gjaltema et al., 1994). More recently, efforts have been made to model biofilm growth using methods that observe structural heterogeneity. Discrete methods (Picioreanu et al., 1997; Wimpenny and Colasanti, 1997) and individual-based modeling (Kreft et al., 1998) have been successfully used to describe macroscopic scale morphology of biofilms based on small-scale interactions. However, the model intrinsic parameters, such as individual cell properties and reactive yields, are usually obtained from planktonic culture experiments in spite of the recognition that planktonic cell culture characteristics do not extrapolate to biofilm systems (Costerton et al., 1994). Therefore, the development of reliable predictive models would greatly benefit from information acquired by time resolved direct observation of biofilms in their naturally hydrated state, as enabled by CLSM (Lawerence et al., 1991).

In a recent report on quantitative analysis of CLSM images of biofilms, Yang et al. (1999) distinguished measurable structural parameters as either textural or areal. According to this classification, "the textural parameters describe the microscale heterogeneity of the image and the areal parameters describe the morphological relationship between the size orientation and shape of surface features". Accordingly, textural parameters are obtained directly from the gray scale image, whereas, areal parameters are measured from a binary image, obtained by segmentation of the gray scale original. That work focused on two-dimensional quantitative analysis. In the present work, the notation introduced by Yang et al. (1999) is extended to three dimensions to identify a fully automated biofilm image analysis procedure. In addition, volume and interfacial area were measured from binary three-dimensional images that, analogously, will be hereby designated as volumetric parameters.

CLSM data are acquired in the form of series of optical cross-sections, i.e. two-dimensional digital images collected in series that scan a given volume at sequential vertical depths. Each digital image is composed by individual discrete elements, pixels, evenly spaced on a rectangular grid. A CLSM scan composed by several two-dimensional images placed at different heights constitutes a three-dimensional grid. The discrete elements that compose the three-dimensional images are called voxels, i.e. volumetric pixels. Segmentation is the process of assigning the image voxels to recognizable components, such as cellular material, phenanthrene crystals or extracellular polymers, producing a binary image for each of them (Castleman, 1996). Direct thresholding has been the most widely used segmentation method, with threshold levels being selected visually by an operator (Yang et al., 1999; Hermanowicz et al., 1995; Kuehn et al., 1998). The widespread use of visual selection reflects the lack of reliable automated methods for determination of threshold levels, and is a serious obstacle to normalization and automation of CLSM for quantitative analysis. An automated method for determination of a threshold level is proposed here. The method introduced in this report, objective threshold selection (OTS), uses a statistical approach similar to that of robust automated threshold selection (RATS) method (Kitler et al., 1985), combined with the analysis of threshold distributions obtained at different positions in the specimen. OTS allows the determination of a median threshold level, more reliable than the one computed using a simple three-dimensional extension of the RATS algorithm,

with the additional advantage of identifying confidence intervals for the estimate.

2. Materials and methods

Two different biofilm systems were used to generate image time series for OTS analysis. The first example consisted of a multiple species denitrifying flow-through cell system. For the second example, a pure culture biofilm using phenanthrene crystals as a carbon source was used.

2.1. Mixed culture denitrifying biofilm growth

A mixed culture biofilm was grown in a flowcell reactor using a mixed species inoculum following methodology described previously (Xavier et al., submitted for publication). A synthetic growth medium was used to provide conditions suitable for denitrification. The inoculum was obtained by repeated enrichment culture obtained from a laboratory reactor operated under denitrifying conditions. Inoculation of the flowcell was performed by re-circulation of the inoculum for 1 h. After this period, the re-circulation loop was opened and regular operation of the system was assumed, i.e., continuous feeding of the flowcell with fresh cell-free medium at a constant flow rate. Liquid culture medium provided suitable conditions for denitrification: 5.0 g l^{-1} KNO_3, 3.0 KCH_3COO, 1.52 g l^{-1} KH_2PO_4, 2.50 g l^{-1} Na_2HPO_4, 0.1 $MgSO_4 \cdot 7H_2O$, 0.25 g l^{-1} $(NH_4)_2SO_4$, 0.05 g l^{-1} $CaCl_2$, 2 mg l^{-1} $FeSO_4 \cdot 7H_2O$, 5 mg/ml EDTA, 1 mg l^{-1} $ZnSO_4 \cdot 7H_2O$, 0.3 mg l^{-1} $MnCl_2 \cdot 4H_2O$, 3 mg l^{-1} H_3BO_3, 2 mg l^{-1} $CoCl_2$ $6H_2O$, 0.1 mg l^{-1} $CuCl_2$, 0.2 mg l^{-1} $NiCl_2$ $6H_2O$ and 0.3 mg l^{-1} $NaMoO_4 \cdot 2H_2O$. The flowcell reactor was build out of two microscope cover slips and costume-made silicone pieces. Silicone tubing was used. For medium inflow, a displacement pump (Watson-Marlow) was used and an approximate linear velocity of 18 flow rate of ml min^{-1} (laminar flow).

2.2. Image acquisition for denitrifying biofilm

The biofilm was stained with 0.1 ml of a 6-μl ml^{-1} aqueous solution of Syto 9 (Molecular Probes, Eugene, OR, USA), injected in the inflow stream 15 min prior to each image acquisition. Medium flow was halted for the 15 min following insertion of the stain to allow the dye to penetrate the biofilm. Syto 9 is a non-specific nucleic acid stain, with excitation/emission maxima of 480–500 nm. The stained biofilm was observed using a Leica TCS-NT laser confocal microscope. The microscope was operated as follows: dual excitation (488 and 568 nm), dual emission (530/30 BP into channel 1 to record Syto 9 fluorescence and 650 LP into channel 2 for propidium iodide fluorescence), 40×1.0 NA oil immersion lens at an Airy disc setting of 0.9. Prior to imaging, PMT settings in both channels were adjusted interactively, such that the RGB overlay (live scan) approximated the fluorescence observed through the oculars with a fluorescein LP filter (i.e., red, green, and yellow cells seen with the eyes were correlated to those imaged in the true-color overlay of the confocal scanner). The flowcell, placed under the CLSM, was not moved during the 40 h of operation in order to always image the same positions. Images where collected at 16, 21, 24, 28, 32, 36 and 40 h after inoculation. The analyzed images represent an area of 250×250 μm^2 with a resolution of 512×512 pixels at 8 bit colordepth (256 gray-values). Each image acquisition event consisted of up to 31 optical sections evenly spaced by a vertical step of approximately 2 μm. Images where stored using the "tiff export" option on the Leica software, which saves each individual cross-section as a standard tiff-format digital picture.

2.3. Phenanthrene degrading biofilm growth

For the growth of *gfp*-labeled *Sphingomonas* sp. A0, a phosphate minimal medium without carbon source was used, with the following composition: 0.88 g l^{-1} Na_2HPO_4, 0.1 g l^{-1} KH_2PO_4, 0.25 $(NH_4)_2SO_4$, 0.05 mg l^{-1} $MgCl_2^* 6H_2O$, 0.03 g l^{-1} $Ca(NO_3)_2^* 4H_2O$, 4 mg l^{-1} Na-EDTA, 1.5 mg l^{-1} $FeCl_2$, 5×10^{-2} mg l^{-1} $MnCl_2^* 4H_2O$, 2×10^{-2} mg l^{-1} $CoCl_2^* 6H_2O$, 1.5×10^{-2} mg l^{-1} $Na_2MoO_4^* 2H_2O$, 1×10^{-2} mg l^{-1} ZnCl2, 2.5×10^{-3} mg l^{-1} LiCl, 2.5×10^{-3} mg l^{-1} $SnCl_2^* 2H_2O$, 5×10^{-3} mg l^{-1} H_3BO_3, 1×10^{-2} mg l^{-1} Kbr, 1×10^{-2} mg l^{-1} KI and 2.5×10^{-3} mg l^{-1} $BaCl_2$. The medium was prepared from concentrated stock solutions. Mineral and buffer stock solutions were autoclaved separately to avoid precipitation of phosphate salts. The internal glass surface of flowcells was covered with phenanthrene crystals (Sigma

Aldricht, St. Louis) by applying first a layer of acetone to fixate the crystals. A medium flow rate of 5 ml min^{-1} was used.

2.4. Image acquisition for phenanthrene degrading biofilm

Gfp-labeled *Sphingomonas* sp. A0 cell fluorescence were detected using excitation at 488 nm and emission at 515 nm (longpass-filter), detailed information is available in Schnell et al., in preparation. Phenanthrene crystals were detected using excitation at 543 nm and emission in the red range (without filter). The fluorescence of cell material and phenanthrene crystals is recorded by the two distinct channels, producing two distinct but superimposing gray scale images. The magnification used was $400\times$ with a $40\times$ oil immersion lens. A 1278×639-μm area within the flow cell was observed, by scanning a four by two set of 319.4×319.4 μm square subsections at a 512×512 pixel resolution using an automated setup (Kuehn et al., 1998). Biofilm thickness was scanned vertically by acquiring optical cross-section at 2-μm steps, up to 64 μm. One set of images was collected each day of the experiment, lasting 7 days following cell inoculation.

2.5. Digital image processing

All image processing was implemented under Matlab programming environment (Matlab 5.3, The Mathworks). Matlab was chosen due to convenience offered for matricial calculus and the capability to emulate parallel processing operations. A software package was developed using Matlab language to implement the algorithms introduced below. Matlab functions and script files are available from the authors upon request. The hardware for computation consisted of Pentium processor-based pc-compatibles running Linux operating system.

2.6. Image segmentation—automated threshold level selection

Original CLSM images, collected with 256 gray-values, must be converted to binary images in order to measure volumetric parameters through a process called image segmentation (Fig. 1). An automated segmentation method of three-dimensional CLSM data sets, named objective threshold selection (OTS), was developed. OTS development was based on the approach introduced by the robust automated threshold selection (RATS) method (Xavier et al., submitted for publication), which states that for a two-dimensional image, A, and in the absence of noise, the optimum threshold (**T**) is given by Eq. (1). In this equation, $p(x,y)$ is the gray-value for the pixel found at coordinates (x,y), and **e** is an edge information matrix for image A. The edge information matrix can be obtained using one of several edge detection algorithms found in the literature (Wilkinson and Schut, 1998).

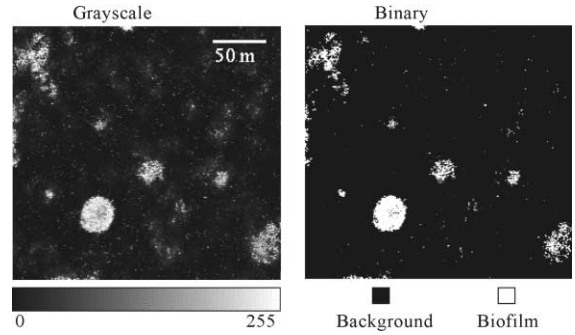

Fig. 1. Gray scale image of an optical cross-section for the denitrifying biofilm, acquired at 28 h and 18 μm from the solid substratum, and corresponding binary image after segmentation. Image segmentation was performed by direct thresholding at gray level 136, which was determined using OTS.

$$\mathbf{T} = \frac{\sum_{A} \mathbf{e}(x,y)\cdot p(x,y)}{\sum_{A} \mathbf{e}(x,y)} \quad (1)$$

RATS can be extended to a three-dimensional data set by replacing $\mathbf{e}(x,y)$ and $p(x,y)$ in Eq. (1) by three-dimensional equivalents, $\mathbf{e}(x,y,z)$ and $p(x,y,z)$. It was demonstrated that, for noisy images, the RATS method yields optimum results when the background and the object each occupy 50% of the image (Xavier et al., submitted for publication). Using the RATS method directly on noisy images where the object is represented in less than 50% of the image will return threshold levels that overestimate the size of the object. Several modifications of the RATS are available for two-dimensional images that compensate low object fractions (Wilkinson and Schut, 1998). Such methods use powers of the edge, $\mathbf{e}^m$ with $m>1$, in Eq. (1), which results in attributing greater statistical weight to higher gradient pixels. The OTS method proposed here is based

on RATS, further developed to handle the particular geometric properties of CLSM biofilm 3D imaging, a description of the algorithm follows.

OTS starts by computing multiple **T**-values, one for each vertical column of the stack of optical sections representing the observed volume, i.e., set of voxels placed on a column which is perpendicular to the solid substratum surface. **T**-values are attributed to each (x,y) position in a surface parallel to the solid substratum (Eq. (2)).

$$\mathbf{T}(x,y) = \frac{\sum_{z=z_0}^{z_n} \mathbf{e}(x,y,z) \cdot p(x,y,z)}{\sum_{z=z_0}^{z_n} \mathbf{e}(x,y,z)} \quad (2)$$

Although matrix $\mathbf{T}(x,y)$ in no way constitutes a vertical projection of the 3D stack, an analogy with maximum projection operations allows better visualization of the concept. Maximum projection operations result in the two-dimensional mapping of the maximum gray-value found along a column of voxels perpendicular to the substratum surface. Similarly, the matrix $\mathbf{T}(x,y)$ is a two-dimensional mapping of the **T**-values determined using RATS along a column of voxels perpendicular to the substratum surface.

The edge matrix, $\mathbf{e}(x,y,z)$, used in OTS is the norm of the numerical gradient of $p(x,y,z)$, which is translated as Eq. (3), obtained by substituting differentials by finite differences.

$$\|\vec{\nabla}_p\| = \left\| \frac{\partial p}{\partial x}\vec{u}_x + \frac{\partial p}{\partial x}\vec{u}_y + \frac{\partial p}{\partial x}\vec{u}_z \right\| \approx \sqrt{\left(\frac{\Delta p}{\Delta x}\right)^2 + \left(\frac{\Delta p}{\Delta y}\right)^2 + \left(\frac{\Delta p}{\Delta z}\right)^2} \equiv \mathbf{e}(x,y,z) \quad (3)$$

The values for $\mathbf{T}(x,y)$ are used to find a cumulative distribution of **T**, $F(\mathbf{T})$, which is then statistically analyzed to determine the median value, $\mathbf{T}_{0.5}$, and an arbitrary confidence level range. A 95% confidence level range, $[\mathbf{T}_{0.025}, \mathbf{T}_{0.975}]$, is suggested for quantitative analysis and will be used below.

Although a main assumption for RATS is that the image is noise-free, the OTS procedure does not have that limitation as it uses the vertical orientation of biofilm geometry to compensate for the effect of noise. The (x,y) positions on the 3D image where no object is found are identified and excluded by OTS, as this is where noise would greatly bias the median **T** to lower values. This is an iterative operation that starts by estimating an initial threshold value, $\mathbf{T}^0$, by applying Eq. (2) to the whole dataset. The value $\mathbf{T}^0$ is then used to reduce the dataset by thresholding a maximum vertical projection image and removing areas in which no object is found. $\mathbf{T}^1$ is found by applying Eq. (2) to the reduced data set. The iterative process continues until a stationary

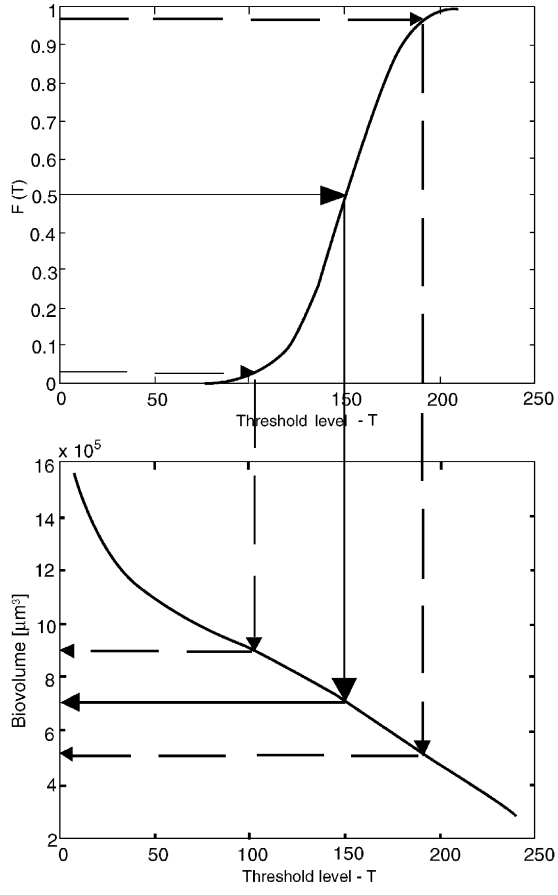

Fig. 2. Measuring biovolumes with confidence 95% levels using the cumulative distribution for threshold levels, $F(\mathbf{T})$. The confidence interval for threshold value is determined from $F(\mathbf{T})$, in the top plot, and propagated to find maximum and minimum values for the confidence interval of the volumetric parameter being estimated, in the bottom plot.

value, $\mathbf{T}^n = \mathbf{T}^{n-1}$, is reached (see Appendix A for MATLAB™ implementation of the OTS algorithm with explanatory annotations).

2.7. Measurement of volumetric parameters

Measurement of volumes is obtained directly from the binary three-dimensional image using a voxel counting algorithm and converting the obtained value using the volume of a voxel in length units (μm^3 in this case). Measurement of interfacial area from binary images is a computationally intensive process, which uses the Matlab™ *isosurface* function to find the coordinates of a set of adjacent polygons that approximates the biofilm interfacial surface. Volumetric parameters are highly dependent of the threshold level used for segmentation and, consequently, determining its values with a statistical confidence level is imperative. For a given volumetric parameter, v, a function of its dependence of the threshold level, $\mathbf{T}$, can be determined, $v(\mathbf{T})$. Then, the cumulative distribution of threshold levels, $F(\mathbf{T})$, identifies confidence intervals that can be propagated for the volumetric parameter being estimated, v (Fig. 2).

3. Volumetric measurements of biofilm CLSM images

3.1. Example 1—Denitrifying biofilm system-time course biovolume measurements with microscope settings optimized for each image acquisition

A mixed species culture was grown under denitrifying conditions in a flowcell reactor for 40 h and monitored using CLSM. Three-dimensional CLSM images of the whole biofilm structure were collected at the same position within the flowcell at 16, 21, 24, 28, 32, 36 and 40 h after inoculation. Microscope settings were tuned prior to each image acquisition to avoid image saturation so that the available 8 bit gray scale is used to its full extent. Each set of images was processed independently for measurement of biovolume and interfacial area, following automated determination of threshold levels.

3.1.1. Image segmentation

The OTS method described above was used for automated determination of threshold levels for image segmentation by direct thresholding. The cumulative distributions of threshold levels, $F(\mathbf{T})$, were used to determine median values, $\mathbf{T}_{0.5}$, and 95% confidence level limits, $\mathbf{T}_{0.025}$ and $\mathbf{T}_{0.975}$ (Table 1).

3.1.2. Biovolume measurements

Biovolume was determined using median threshold levels, $\mathbf{T}_{0.5}$. The measurement error associated with the segmentation is reflected in the confidence intervals for biovolume. Biovolume measurement, V, was observed to be highly dependent of the threshold level used for segmentation (Fig. 3A). The 95% confidence interval is determined by finding biovolume maximum and minimum within $[\mathbf{T}_{0.025}, \mathbf{T}_{0.975}]$. Biovolume was observed to grow exponentially during the period studied, with a growth rate of $\mu_V = 0.09 \pm 0.03\ h^{-1}$. Therefore, the volumetric measurements were obtained for a sixfold increase in biovolume. It is noteworthy that the absolute error was found to be independent of the volume measured (Fig. 3B).

3.1.3. Biofilm interfacial area measurements

Median values for the interfacial area were determined from thresholding at $\mathbf{T}_{0.5}$. The measurement error associated with the segmentation procedure is determined by plotting interfacial area as a function of the threshold level (Fig. 5A). The interfacial area, function of the $\mathbf{T}$-level, $A(\mathbf{T})$, is not always monotonous within the $[\mathbf{T}_{0.025}, \mathbf{T}_{0.975}]$, as opposed to $V(\mathbf{T})$ which typically decreases with $\mathbf{T}$. Interfacial area for the biofilm was found to increase in the 40 h of the experiment with a specific growth rate of $\mu_A = 0.5 \pm 0.03\ h^{-1}$ (Fig. 4B).

Table 1
Automated threshold levels obtained by OTS of CLSM images from the denitrifying flow cell (example 1)
The median value, $\mathbf{T}_{0.5}$, and 95% confidence levels were extracted from the cumulative distribution $F(\mathbf{T})$, such as the one represented in Fig. 3. Each time sample of images was analyzed independently. The threshold values listed in this table delimit the confidence intervals propagated in Figs. 4A and 5A.

Time (h)	$\mathbf{T}_{0.025}$	$\mathbf{T}_{0.5}$	$\mathbf{T}_{0.975}$
16	173	47	18
21	188	126	70
24	196	134	79
28	185	136	88
32	193	151	104
36	194	151	104
40	192	151	102

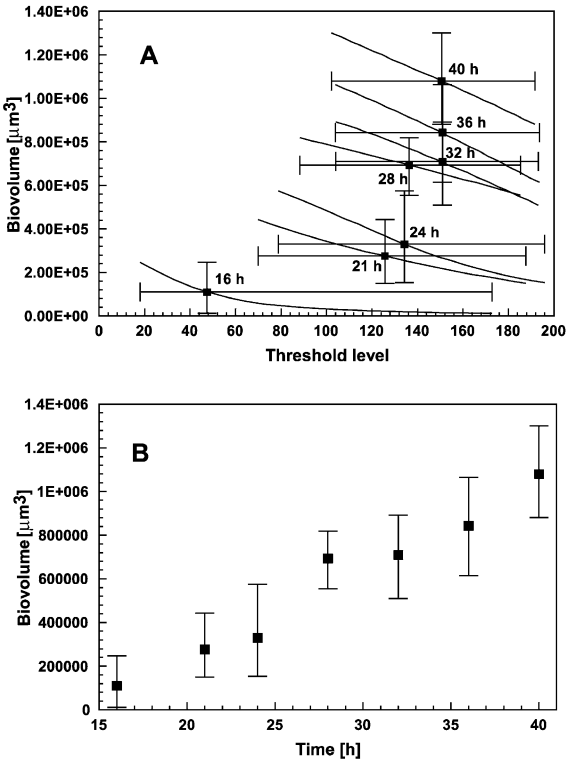

Fig. 3. (A) Distribution of biovolume estimate as a function of the threshold value used for image segmentation. Median values for biovolume are represented by the black symbols (■) on each curve corresponding to a time image acquisition. The lines represent the dependency of volume on segmentation threshold value within the 95% confidence interval for both parameters (notice that the vertical and horizontal errors intersect at the ends of the line). Confidence interval for biovolume was propagated from the threshold distribution (Table 1), as shown in Fig. 2. (B) Biovolume measurements for the denitrifying biofilm from 21 to 40 h after inoculation. Error bars represent 95% confidence intervals obtained from the dependency on segmentation threshold plotted in (A).

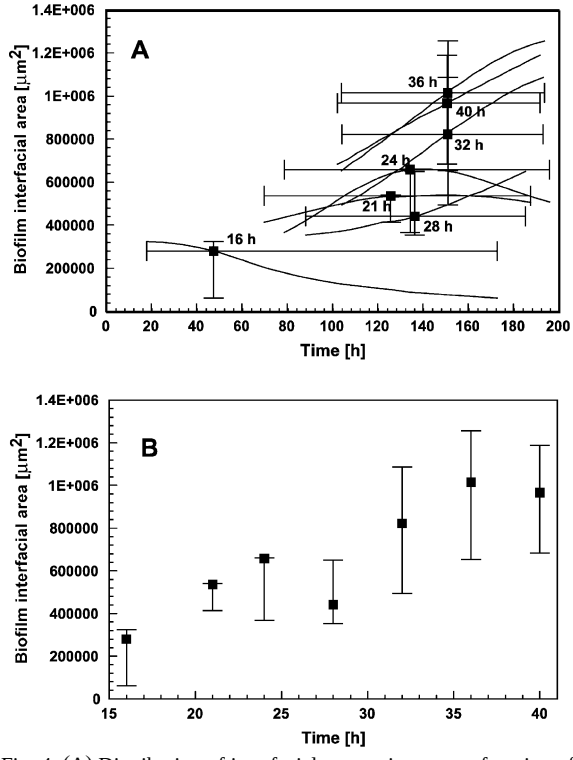

Fig. 4. (A) Distribution of interfacial area estimate as a function of the threshold value used for image segmentation. Median values for interfacial area are represented by the black symbols (■) on each curve corresponding to a time image acquisition. The lines represent the dependency of volume on segmentation threshold value within the 95% confidence interval for both parameters (notice that the vertical and horizontal errors intersect at the ends of the line). Confidence intervals for interfacial area were propagated from the threshold distribution (Table 1), as shown in Fig. 2 for volume estimation. (B) Interfacial area measurements for the denitrifying biofilm between 21 and 40 h after inoculation. Error bars represent 95% confidence intervals obtained from the dependency on segmentation threshold plotted in (A).

3.2. Example 2—Biofilm grown on phenanthrene crystal-time course biovolume measurements with fixed microscope optical settings

The growth of a monospecies biofilm of a phenanthrene degrading *Sphingomonas* sp. strain in the presence of phenanthrene crystals as the sole carbon source was monitored for 7 days. The biofilm was observed through *gfp* expression of the *Sphingomonas* strain. Phenanthrene crystal particles, placed on the surface of the solid substratum where biofilm growth took place, were observed by autofluorescence on red wavelengths. CLSM images of both the green and red color channels were collected once a day for the duration of the experiment. A four by two set of adjacent three-dimensional image stacks was collected for each color channel per day, following the automated image acquisition procedure described in Materials and methods. Unlike the previous example, here the optical settings of the microscope were optimized for the conditions of the first imaging, on day 1, and maintained throughout the duration of the experiment. The image stacks obtained from the two channels were processed inde-

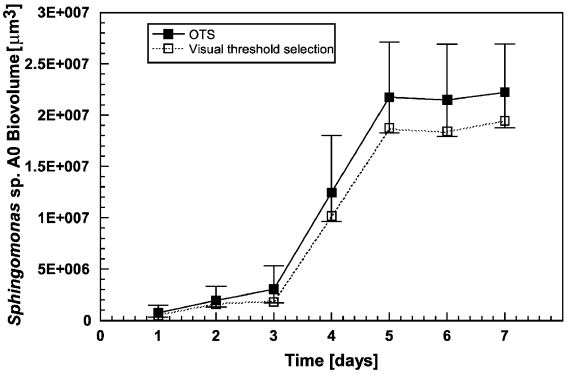

Fig. 5. Biovolume measurements for the monospecies phenanthrene degrading *Sphingomonas* sp. A0 biofilm. Comparison between biovolume estimates obtained from automated OTS segmentation and visual threshold selection by an experienced operator. Confidence intervals for the OTS estimates correspond to a 95% probability level identified by propagation of threshold values, as described in Fig. 2.

pendently for the determination of threshold levels and biovolume measurements. Only the latter are presented here (Fig. 5). It can be seen in this figure that, unlike the previous example where error intervals were observed to stay relatively constant, they increase as biofilm growth progresses. The fact that settings were selected at low quantities of biomass causes images collected at later growth stages to show increasing brightness saturation. In this plot, biovolume estimates using OTS were also compared with values obtained using segmentation by visual threshold selection by an experienced operator. It was observed that the two were similar, with values obtained manually falling within the 95% confidence intervals of the automated estimates. Volume estimates based on manual segmentation do not generate objective confidence intervals, as that would require repeated visual segmentation by several experienced operators, not a practical proposition.

4. Discussion

The OTS method presented was evaluated by processing CLSM images from two different biofilm systems, where time-resolved data was generated describing biofilm morphogenesis. The first example consisted of a mixed species biofilm grown under denitrifying conditions for 40 h. Three-dimensional images were acquired at seven time points, ranging from 16 to 40 h after inoculation. For this first case, microscope settings were reset by the operator prior to each image acquisition. In the second system, growth from a mono-species biofilm from phenanthrene degrading *Sphingomonas* sp. strain A0 was imaged in the presence of phenanthrene crystals as the sole carbon source. The CLSM images were acquired once every day, up to 7 days after inoculation. In this case, microscope settings were selected for the first image acquisition, right after flowcell inoculation, and maintained throughout the entire experiment. The fact that settings were chosen when the biofilm showed least fluorescence caused images collected at later times, when a larger amount of biomass was present, to show more brightness saturation, which was reflected by an increase in the estimated error (Fig. 5). This explains why the estimates with the steadier confidence intervals (Figs. 3 and 4) are achieved by resetting image acquisition parameters for each image acquisition in order to minimize the proportion of saturated voxels. Nevertheless, the median estimates were observed to be relatively insensitive to sub-optimal microscope settings, which are mostly reflected in the broader confidence interval. Similarly, degradation in the quality of the image in general was observed to have a noticeable effect in the confidence interval of volumetric estimates. The robustness of the median value is based on the even distribution of **T**-values. There is, however, a cause of loss in image quality that will be not apparent in the analysis of confidence distributions—insufficient light penetration to the furthest optical sections of thick specimens. This can be understood by noting that a single threshold value is obtained for each vertical voxel column and dim fluorescence in the furthest positions due to insufficient light penetration or thinning of cellular material would be undistinguishable. Nevertheless, this is a circumstance that the CLSM operator would promptly identify.

5. Conclusion

The image analysis procedure proposed here, named optical threshold selection (OTS), automates the determination of threshold levels for 3D-image segmentation. This method was applied to biofilm images obtained by confocal laser scanning mi-

croscopy and was shown to reproduce values obtained manually by experienced operators. In addition to the savings in time and manpower, the proposed method is amenable to a statistical analysis to extract confidence intervals, a feature not practical with the current manual selection procedure. The algorithm developed is based on earlier work on robust threshold selection (RATS) in 2D images, which is generalized for three and higher dimensions —a time series of 3D images is a 4D object. The determination of confidence intervals follows a non-parametric approach, not making any assumption regarding threshold distribution in the image. The identification of confidence intervals for volume and area estimates is obtained by propagation of the observed distribution. OTS enables segmentation of images of any dimension that is both independent from acquisition settings and fully automated.

Appendix A. Implementation of the algorithm on MATLAB™—code and example of application

The algorithm for the OTS was fully implemented on MATLAB™ 5.3 and is part of the CLSM toolbox, developed by the authors, which is available freely upon request, and may also be downloaded from http://www.itqb.unl.pt:1111/~jxavier/clsmtoolbox/clsmtoolbox.zip.

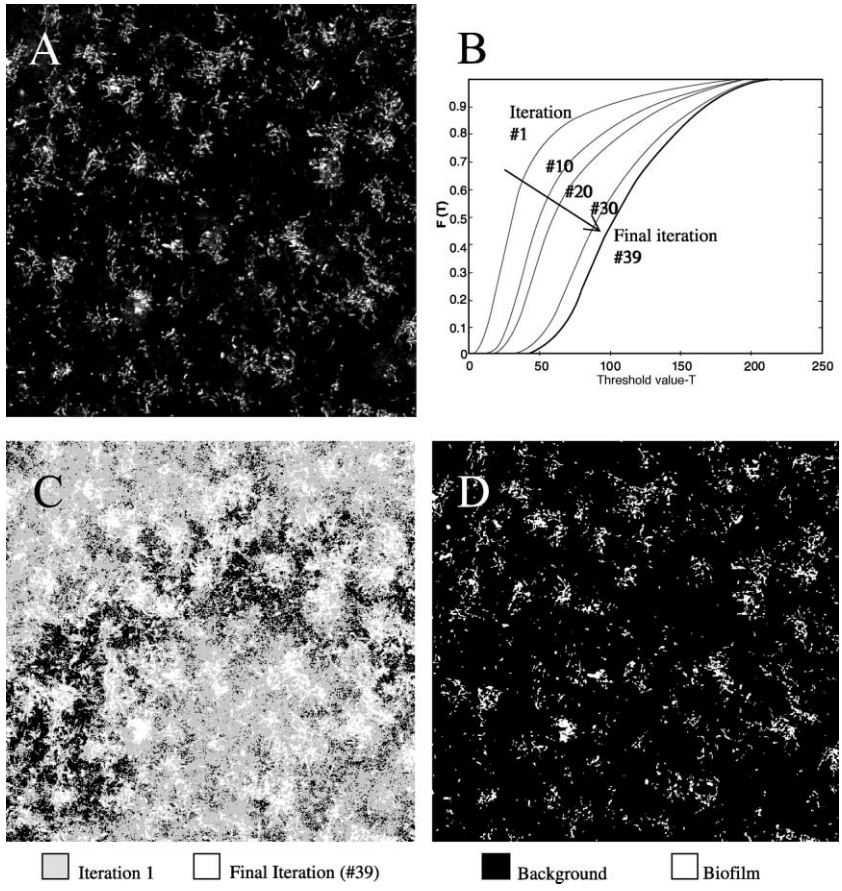

Fig. 6. Illustration of the iterative process used to compensate for the effect of noise on the determined threshold (accompanying code listed). (A) Original optical cross-section (gray scale) of a biofilm, part of a 3D stack. (B) Changes on the cumulative distribution of threshold values, $F(\mathbf{T})$ computed following Eq. (2), throughout the iterations. $F(\mathbf{T})$ is used to estimate a mean threshold level, which is then used on a projection images of the biofilm. Only the area delimited by the maximum projection in the previous iteration is then used to determine $F(\mathbf{T})$ on the following iteration. (C) Area of the substratum covered by biofilm according to the first and final iterations. (D) Image after segmentation by direct thresholding with the final mean **T**-value.

The CLSM toolbox contains implementation of several image processing operations, as well as functions for loading CLSM stacks onto the MATLAB workspace. In the CLSM toolbox package, there is also a detailed reference manual in html format, which includes instructions for installation under Microsoft Windows™. An online version of the HTML manual for the CLSM toolbox may also be found at http://www.itqb.unl.pt:1111/~jxavier/clsmtoolbox/clsm_toolbox_help/index.html.

The integral transcription of the MATLAB™-code for the pivotal function *iterative_cdvr_Ft*, which is central to the OTS method, is listed below. The code is annotated with comments that refer to the illustration of its processing in Fig. 6.

```
function [Ft,T50,T975,T025,nit]=iterative_cdvr_Ft(im,Dz)

% [Ft,T50,T975,T025,nit]=iterative_cdvr_Ft(im,Dz) – Cumulative
distribution of vertical RATS
% and threshold value (with 95% confidence level) following the OTS
(Objective Threshold Selection) method
%
% Computes RATS cumulative distribution with iterations to
% increase weight of colonized areas.
%
%   Input arguments:
%         im   -  three-dimensional matrix with stack of CLSM images (Fig. 6A)
%         Dz   -  ratio of vertical step (distance between optical cross
sections)
%         by the pixel size of optical cross section images
%
%   Output arguments:
%         Ft       -  cumulative distribution of vertically oriented thresholds
%         T50    -  mean threshold value, such that F(T50) = 50%
%         T975  -  upper limit of the threshold value 95% confidence level,
such that F(T975) = 97.5%
%         T025  -  lower limit of the threshold value 95% confidence level,
such that F(T025) =  2.5%
%
% João Xavier (jxavier@itqb.unl.pt) - September 2000

if (nargout~=5),
     error('iterative_cdvr_Ft requires 5 output arguments ');
end;

[m,n,q]=size(im);
im=double(im);


% defining x,y and z oriented gradient kernels
Bx=zeros(3,3,3); Bx(2,1,2)=1; Bx(2,2,2)=-1;
By=zeros(3,3,3); By(1,2,2)=1; By(2,2,2)=-1;
Bz=zeros(3,3,3); Bz(2,2,1)=1/Dz; Bz(2,2,2)=-1/Dz;

% computing edge information
imedge=sqrt(ofilter3(Bx,im).^2+ofilter3(By,im).^2+ofilter3(Bz,im).^2);
```

```
% removing border pixels that have distorted information:
deno_nb=imedge(2:m-1,2:n-1,2:q-1);
num_nb=deno_nb.*im(2:m-1,2:n-1,2:q-1);

[M,N,Q]=size(deno_nb);

deno_a=reshape(sum((reshape(deno_nb,M*N,Q))'),M,N);
num_a=reshape(sum((reshape(num_nb,M*N,Q))'),M,N);

deno=deno_a(find(and(deno_a~=0,num_a~=0)));
num=num_a(find(and(deno_a~=0,num_a~=0)));

% Finding first estimate of mean thresholding Tm
RATSs=sort(num./deno);
Tm=RATSs(round(length(RATSs)./2));

% iterative cicle (see Fig. 6B and 6C):
imtp=reshape(sum((reshape(im(2:m-1,2:n-1,2:q-1)>Tm,M*N,Q))'),M,N);
%projection of thresholded image
cfrac=sum(imtp(:)>0)./(M*N);
nit=1;
Tma=-1; %initial value for the cycle check
while Tm~=Tma,
    Tma=Tm;
    nit=nit+1;
    deno2=deno_a(find((deno_a~=0).*(num_a~=0).*(imtp>0)));
    num2=num_a(find((deno_a~=0).*(num_a~=0).*(imtp>0)));
    RATS2=num2./deno2;
    RATSs=sort(RATS2);
    Tm=RATSs(round(length(RATSs)./2));
    imtp2=reshape(sum((reshape(im(2:m-1,2:n-1,2:q-1)>Tm,M*N,Q))'),M,N);
    cfrac=sum(and(imtp2(:)>0,imtp(:)>0))./sum(imtp(:)>0);
    imtp=imtp2;
end;

% Attribute values for output arguments
Ft=RATSs;
T50=RATSs(round(length(RATSs)*.5));
T975=RATSs(round(length(RATSs)*.975));
T025=RATSs(round(length(RATSs)*.025));
```

References

Caldeira, M., Heald, S.C., Carvalho, M.F., Vasconcelos, I., Bull, A.T., Castro, P.M., 1999. 4-Chlorophenol degradation by a bacterial consortium: development of a granular activated carbon biofilm reactor. Appl. Microbiol. Biotechnol. 52 (5), 722–729.

Castleman, K.R., 1996. Digital Image Processing. Prentice-Hall, London.

Christensen, B.B., Sternberg, C., Andersen, J.B., Eberl, L., Moller, S., Givskov, M., Molin, S., 1998. Establishment of new genetic traits in a microbial biofilm community. Appl. Environ. Microbiol. 64 (6), 2247–2255.

Costerton, J.W., Lewadowski, Z., DeBeer, D., Caldwell, D., Ko-

rver, D., James, G., 1994. Biofilms: the customized microniche. J. Bacteriol. 176 (8), 2137–2142.

DeBeer, D., Stoodley, P., Roe, F., Lewadowski, Z., 1994. Effects of biofilm structures on oxygen distribution and mass transport. Biotechnol. Bioeng. 43, 1131–1138.

Gjaltema, A., Arts, P.A.M., van Loosdrecht, M.C.M., Kuenen, J.G., Heijnen, J.J., 1994. Heterogeneity of biofilms in rotating annular reactors: occurrence, structure, and consequences. Biotechnol. Bioeng. 44 (2), 194–204.

Hermanowicz, S.W., Schindler, U., Wilderer, P., 1995. Fractal structure of biofilms: new tools for investigation of morphology. Water Sci. Technol. 32 (8), 99–105.

Jayaraman, A., Sun, A.K., Wood, T.K., 1998. Characterization of axenic *Pseudomonas fragi* and *Escherichia coli* biofilms that inhibit corrosion of SAE 1018 steel. J. Appl. Microbiol. 84 (4), 485–492.

Kitler, J., Illingworth, J., Föglein, J., 1985. Threshold selection based on a simple image statistic. Comput. Vision Graph Image Process. 30, 125–147.

Kreft, J.U., Booth, G., Wimpenny, J.W.T., 1998. BacSim: a simulator for individual-based modelling of bacterial colony growth. Microbiology 144, 3275–3287.

Kuehn, M., Hausner, M., Bungartz, H.-J., Wagner, M., Wilderer, P.A., Wuertz, S., 1998. Automated confocal laser scanning microscopy and semi-automated image processing for analysis of biofilms. Appl. Environ. Microbiol. 64 (11), 4115–4127.

Lawerence, J.R., Korber, D.R., Hoyle, B.D., Costerton, J.W., Caldwell, D.E., 1991. Optical sectioning of microbial biofilms. J. Bacteriol. 173, 6558–6567.

Lawerence, J.R., Wolfaardt, G.M., Korber, D.R., 1994. Determination of diffusion coefficients in biofilms by confocal laser microscopy. Appl. Environ. Microbiol. 60 (4), 1166–1173.

Moller, S., Sternberg, C., Andersen, J.B., Christensen, B.B., Ramos, J.L., Givskov, M., Molin, S., 1998. In situ gene expression in mixed-culture biofilms: evidence of metabolic interactions between community members. Appl. Environ. Microbiol. 64 (2), 721–732.

Picioreanu, C., van Loosdrecht, M.C.M., Heijnen, J., 1997. Mathematical modeling of biofilm structure with a hybrid differential-discrete cellular automaton approach. Biotechnol. Bioeng. 58 (1), 101–116.

Schnell, A., Wattiau, P., Springael, D., Wuertz, S., in preparation. Dual in situ detection of PAH-degrading *Sphingomonas* sp. LB126 based on gfp labeling and a highly specific 16S rRNA-targeted oligonucleotide.

Sternberg, C., Christensen, B.B., Johansen, T., Toftgaard Nielsen, A., Andersen, J.B., Givskov, M., Molin, S., 1999. Distribution of bacterial growth activity in flow-chamber biofilms. Appl. Environ. Microbiol. 65 (9), 4108–4117.

Stoodley, P., Lewandowski, Z., Boyle, J.D., Lappin-Scott, H.M., 1998. Oscillation characteristics of biofilm streamers in turbulent flowing water as related to drag and pressure drop. Biotechnol. Bioeng. 57 (5), 536–544.

Wilkinson, M.H.F., Schut, F., 1998. Digital Analysis of Microbes. Wiley, Chichester.

Wimpenny, J.T., Colasanti, R., 1997. A unifying hypothesis for the structure of microbial biofilms based on cellular automaton models. FEMS Microbiol. Ecol. 22, 1–16.

Wood, P., Caldwell, D.E., Evans, E., Jones, M., Korber, D.R., Wolfhaardt, G.M., Wilson, M., Gilbert, P., 1998. Surface-catalysed disinfection of thick *Pseudomonas aeruginosa* biofilms. J. Appl. Microbiol. 84 (6), 1092–1098.

Xavier, J.B., Malhó, R., Reis, A.M., Palmer, R., White, D.C., Almeida, J.S., submitted for publication. Biofilm morphogenesis is self organizing.

Yang, X., Beyenal, H., Harkin, G., Lewadowski, Z., 1999. Quantifying biofilm structure using image analysis. J. Microbiol. Methods 39, 109–119.