PERGAMON
www.elsevier.com/locate/watres

PII: S0043-1354(00)00361-4

Wat. Res. Vol. 35, No. 5, pp. 1149–1158, 2001
© 2001 Elsevier Science Ltd. All rights reserved
Printed in Great Britain
0043-1354/01/$ - see front matter

EVALUATION OF BIOFILM IMAGE THRESHOLDING METHODS

XINMIN YANG[1], HALUK BEYENAL[1], GARY HARKIN[2] and ZBIGNIEW LEWANDOWSKI[1,3]*

[1]Center for Biofilm Engineering, Montana State University, Room 366 EPS, P.O. Box 173980, Bozeman, MT 59717-3980, USA; [2]Computer Science Department, Montana State University, Bozeman, MT 59717-3980, USA and [3]Department of Civil Engineering, Montana State University, Bozeman, MT 59717-3980, USA

(First received 1 January 2000; accepted in revised form 2 June 2000)

Abstract—To evaluate biomass distribution in heterogeneous biofilms from their microscope images, it is often necessary to perform image thresholding by converting the gray-scale images to binary images consisting of a foreground of biomass material and a background of interstitial space. The selection of the gray-scale intensity used for thresholding is arbitrary but under the control of the operator, which may produce unacceptable levels of variability among operators. The quality of numerical information extracted from the images is diminished by such variability, and it is desirable to find a method that improves the reproducibility of thresholding operations. Automatic methods of thresholding provide this reproducibility, but often at the expense of accuracy, as they consistently set thresholds that differ significantly from what human operators would choose. The performance of five automatic image thresholding algorithms was tested in this study: (1) local entropy; (2) joint entropy; (3) relative entropy; (4) Renyi's entropy; and (5) iterative selection. Only the iterative selection method was satisfactory in that it was consistently setting the threshold level near that set manually. The extraction of feature information from biofilm images benefits from automatic thresholding and can be extended to other fields, such as medical imaging. © 2001 Elsevier Science Ltd. All rights reserved

Key words—biofilm, biofilm structure, imaging, image thresholding, segmentation

NOMENCLATURE

b_1	binary gray level of background (0)
b_o	binary gray level of foreground (1)
$f(m,n)$	the gray-scale value of the pixel located at point position (m, n)
$f_t(m,n)$	the pixel values of a thresholded biofilm image
G	set of gray-scale levels
$H(t)^{\alpha}_{\Lambda 1}$	background Renyi's entropy
$H(t)^{\alpha}_{\Lambda 2}$	foreground Renyi's entropy
$H_{B\to B}(t)$	background-to-background entropy
$H_{F\to F}(t)$	foreground-to-foreground entropy
H_{joint}	joint entropy
H_{local}	local entropy
$H_{relative}$	relative entropy
M	horizontal size of the matrix, integer
MSSRR	mean of the sum of squares of relative residuals
N	vertical size of the matrix, integer
N_G	maximum gray-scale value (255)
N_m	total number of images being tested
$P^{(x)}$	probability distributions in quadrant X of a binary image
p_i	probability of a certain gray level at point i
$P_x(t)$	probability of gray level change in quadrant X
P^X_{ij}	probabilities of gray level change in quadrant X of co-occurrence matrix
$q_x(t)$	probability distributions in quadrant X of a binary image
t	threshold
t^*_c	optimal threshold
t_i*	gray-scale level which maximizes the sum of Renyi's entropy
t_{ij}	(i, j)th entry of co-occurrence matrix
va_i	average manual threshold
v_i	threshold calculated from automatic method

Greek letters

β	variable used to calculate Renyi's entropy
Λ_1	histogram probability of object
Λ_2	histogram probability of background
α	real positive number $\neq 1$
ω	variable used to calculate Renyi's entropy

INTRODUCTION

The conceptual image of biofilms has recently evolved from displaying microorganisms uniformly distributed in the matrix of extracellular biopolymers to structurally heterogeneous biofilms displaying microorganisms aggregated in microcolonies

*Author to whom all correspondence should be addressed. Center for Biofilm Engineering, P.O. Box 173980, Room 366 EPS, Montana State University, Bozeman, MT 59717-3980, USA. Tel.: 1-406-994-5915; fax: 1-406-994-6098; e-mail: zl@erc.montana.edu

(Costerton *et al.*, 1994; Bishop and Rittmann, 1995; Wolfaardt *et al.*, 1994; Massol-Deya, 1994; Keevil and Walker, 1992). Although conceptual models of heterogeneous biofilms produced by various research groups differ in details, they all declare that the biofilm matrix is porous, thus allowing nutrients to reach deep layers of the biofilm. The acceptance of the new conceptual model of biofilm structure has profound influence on the understanding of the intra-biofilm mass transport dynamics. Heterogeneous biofilms, instead of being gelatinous layers of uniform density offering a uniform distribution of mass transport resistance, are now characterized by a three-dimensional distribution of biomass density and, consequently, by the spatial distribution of mass transport resistance (Rittmann *et al.*, 1999). As a consequence of biofilm heterogeneity, the nutrients at different locations penetrate the biofilm with different rates influencing the spatial distribution of microbial activity. To characterize the distribution of microbial activity in such biofilms it is beneficial to quantify the spatial distribution of biomass (Noguera *et al.*, 1999). Expressing the spatial distribution of the biomass in biofilms—biofilm heterogeneity—on a numerical scale allows correlation with the spatial distribution of other parameters; for example, the local mass transport coefficient (Yang and Lewandowski, 1995) or effective diffusivity (Beyenal *et al.*, 1998), therefore providing an insight into the dynamics of the intra-biofilm mass transport.

Our laboratory has developed procedures to monitor spatial and temporal changes in distribution of biomass in biofilms using optical and laser (confocal) microscopy combined with image analysis (Lewandowski *et al.*, 1999; Yang *et al.*, 2000). We have also developed a software that can extract several parameters quantifying the distribution of the biomass from biofilm images, such as areal porosity, diffusional distance, vertical and horizontal sizes of the cell clusters, total perimeters of the cell clusters, fractal dimension and textural entropy (Yang *et al.*, 2000). Extracting numerical parameters from the images involves several steps. One of those steps is related to the image manipulation procedure called thresholding. The images taken by a camera are gray-scale images, while the images processed to extract areal porosity and several other parameters are binary, having two colors only. Images of biofilms are captured with 256 gray-scale levels and thresholding is used to segment them into two parts, biomass and interstitial space. The operator arbitrarily selects one of the 256 gray-scale levels as a threshold. Computer software then segments the image(s) into two parts with all gray levels less than or equal to the threshold becoming a background color and rest of all gray levels becoming a foreground color.

To extract statistically meaningful morphological parameters from a series of images, the thresholding has to be reproducible. There are no rules for setting the threshold, so choice of gray-scale intensity selected to segment the images depends then on the operator's understanding of the image content and the desired relationship between the two segments in the final image. Although it is impossible to specify precisely how to select the threshold, the parameters calculated from thresholded images depend on the selected gray-scale level. The operator uses his or her best judgement and sets the gray-scale level in such a way that the binary image appears to capture the essence of the biofilm structure. Variability between operators in choosing the threshold adversely affects the measurements obtained from the binary image. Figure 1 shows the impact that threshold has on the calculation of biofilm areal porosity (the calculations were performed as described by Yang *et al.*, 2000). Different operator preferences for threshold doubled the estimate of biofilm areal porosity in some cases.

Since there is no absolute measure of correctness for the threshold, we determined that a reasonable measure of effectiveness is that an acceptable automatic thresholding method should approach the average value determined by a panel of human experts. We assumed that an automatic procedure that has sufficient precision will yield more reproducible results

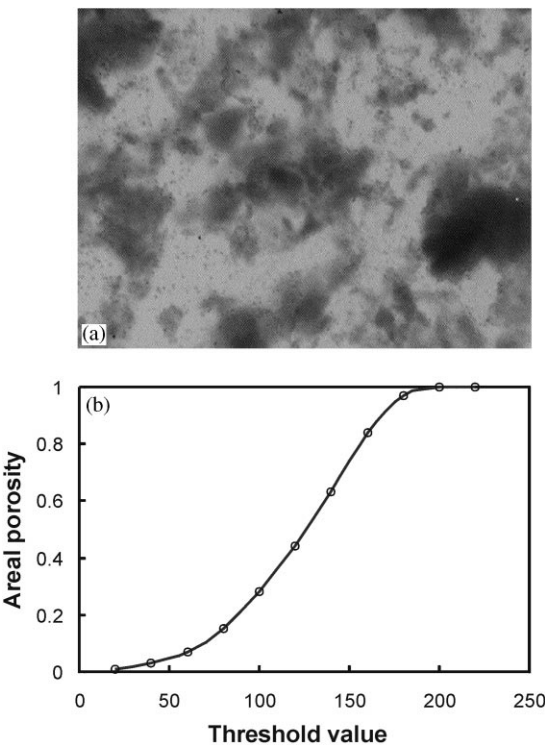


Fig. 1. (a) A gray-scale image of a biofilm. Dark and white areas correspond to biomass and interstitial space, respectively. (b) Biofilm areal porosity calculated for different gray-scale thresholds. The areal porosity was calculated as the ratio of cluster area to total area (see Yang *et al.*, 2000). The image size was $1568 \times 1176\,\mu m$. The porosity was 0.3 and 0.6 for the gray-scale thresholds of 100 and 140, respectively.

than human operators. To evaluate the efficiency of selected automatic image thresholding procedures, we tested: (1) the variability among thresholds set manually by different operators, (2) the variability of thresholds set automatically using known thresholding algorithms, (3) the agreement between the thresholds set manually and automatically. The goal of the work was to determine the relative effectiveness of different automatic thresholding procedures and if any met the acceptability standards.

To our knowledge, image thresholding algorithms have not been tested on biofilm images. Most thresholding algorithms described in the literature were based on measuring boundaries and edges, and were tested on objects having well-defined boundaries such as microbes or fluorescent particles (see review paper by Wilkinson 1998). We tried using these boundary and edge-detecting algorithms for biofilm images. However, the cell clusters in biofilms do not have clearly marked boundaries and the applicability of these algorithms was poor. So, we focused on techniques that used entropy and histograms, and chose five methods of automatic thresholding available in the computer science literature: (1) local entropy (Pal and Pal, 1989); (2) joint entropy (Pal and Pal, 1989); (3) relative entropy (Chang *et al*., 1994); (4) Renyi's entropy (Sahoo *et al*., 1997); and (5) iterative selection (Ridler and Calvard, 1978).

The images of biofilms were generated by reflected light microscopy, transmission light microscopy, and confocal scanning laser microscopy (CSLM). To generate a reference set, 10 researchers from the Center for Biofilm Engineering at Montana State University, Bozeman, manually thresholded the same images. The results of different automatic thresholding procedures were compared with each other and with the results of manual thresholding.

PRINCIPLES OF IMAGE THRESHOLDING

We view thresholding as segmenting an image into foreground and background pixels (Chang *et al*., 1995; Lee and Chung, 1990) based on the choice of a specific gray-scale value. The foreground is biomass and the background is interstitial space. Thresholding methods can be multilevel, but that case is not considered here. Automatic image thresholding methods seek to optimize a threshold-dependent criterion function that is related to the properties of the image (Brink, 1992). The issues of importance are the choice of the criterion function and the method for converging on the optimal value.

Every biofilm image has a histogram which shows the variation of the total number of gray-scale values with gray-scale level in the image. If the object is distinct from the background, the histogram of the image is typically bimodal, with the foreground clustered in one part of the histogram and the background in another. Based on this feature, a threshold can be chosen at the valley of the gray-level histogram.

Many biofilm images are unimodal, which means there is only one peak in the histogram and no obvious threshold choice. We explored two approaches to image thresholding: (1) co-occurrence matrix methods and (2) histogram-based methods. Two-dimensional entropy methods based on the co-occurrence matrix (Haralick *et al*., 1973) are used as measures of thresholding performance. The co-occurrence-based entropy types employed are local entropy (Pal and Pal, 1989), joint entropy (Pal and Pal, 1989) and relative entropy (Chang *et al*., 1994). The histogram-based thresholding methods used are Renyi's entropy (Chang *et al*., 1994) and iterative selection (Ridler and Calvard, 1978).

Thresholding a digital image

If $f(m,n)$ is the gray-scale value of the pixel located at point (m,n), the digital image can be represented as

$$\{f\,(m,n)|m \in \{0,1,2,\ldots,M-1\},\ n \in \{0,1,2,\ldots,N-1\}, f(m,n) \in G\} \quad (1)$$

where the size of the image is $M \times N$ and the set of gray-scale values is G and defined as

$$G = 0,..N_G - 1 \quad (2)$$

Let t be a threshold and $B = \{b_0, b_1\}$ be a pair of binary gray levels with $(b_0, b_1 \in G)$. The result of thresholding the image at gray level, t, is a binary image which is represented by $f_t(m,n)$

$$f_t(m,n) = \begin{cases} b_0 & \text{if } f(m,n) \leq t \\ b_1 & \text{if } f(m,n) > t \end{cases} \quad (3)$$

For a thresholded biofilm image $b_0 = 1$ (black, foreground) and $b_1 = 1$ (white, background).

CO-OCCURRENCE MATRIX BASED METHODS

Calculating co-occurrence matrices

The co-occurrence matrix is a transition matrix generated by changes in pixel intensities (Chang *et al*., 1994), which describes all intensity transitions from gray-scale value i to gray-scale value j. It is a two-dimensional representation of the probability of gray-scale level changes in an image (Yang *et al*., 2000). The element of the (i, j)th entry of co-occurrence matrix is defined as

$$t_{ij} = \sum_{l=0}^{M-1} \sum_{k=0}^{N-1} \delta(l,k) \quad (4)$$

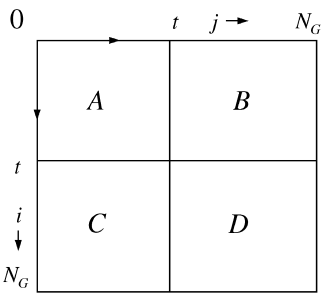
where

$$\delta(l,k) = \begin{cases} 1 & f(l,k) = i \text{ and } f(l,k+1) = j \text{ or } f(l,k) \\ & \quad = i \text{ and } f(l+1,k) = j \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

A simple example of calculating the co-occurrence matrix of a digital biofilm image is given in Yang *et al.* (2000).

Local entropy method

The local entropy method (Pal and Pal, 1989) is based on Shannon's entropy and utilizes the elements of the co-occurrence matrix (Sahoo *et al.*, 1997). Let $t \in G$ be the threshold so that the co-occurrence matrix is partitioned into four quadrants, (see below) A, B, C, and D as shown (N_G is the maximum gray-scale value 255). These correspond to the transition from background-to-background ($B \to B$), background-to-foreground ($B \to F$), foreground-to-background ($F \to B$), and foreground-to-foreground ($F \to F$). The local entropy is defined by the quadrants A and D ($B \to B$ and $F \to F$).



According to the algorithm proposed by Chang *et al.* (1994), the probability of a pixel being in each quadrant is defined as

$$P_A(t) = \sum_{i=0}^{t} \sum_{j=0}^{t} p_{ij} \tag{6}$$

$$P_B(t) = \sum_{i=0}^{t} \sum_{j=t+1}^{N_G} p_{ij} \tag{7}$$

$$P_C(t) = \sum_{i=t+1}^{N_G} \sum_{j=0}^{t} p_{ij} \tag{8}$$

$$P_D(t) = \sum_{i=t+1}^{N_G} \sum_{j=t+1}^{N_G} p_{ij} \tag{9}$$

The probability of an $i \to j$ transition can be normalized by dividing by the probability of being in the $i \to j$ quadrant

$$p_{ij}^A = \frac{p_{ij}}{P_A(t)},$$

$$p_{ij}^B = \frac{p_{ij}}{P_B(t)}, \quad p_{ij}^C = \frac{p_{ij}}{P_C(t)}, \quad p_{ij}^D = \frac{p_{ij}}{P_D(t)} \tag{10}$$

The resulting probabilities are called the transition probability distribution (Chang *et al.*, 1994). The local entropy method uses the spatial correlation in the image as the criteria for selecting the optimal threshold, by attempting to maximize the probability that similar pixels are grouped together. It is based on the probability distribution for foreground-to-foreground and background-to-background transitions in the co-occurrence matrix:

Background-to-background entropy: $H_{B \to B}(t)$

$$= -\frac{1}{2} \sum_{i=0}^{t} \sum_{j=0}^{t} p_{ij}^A \log p_{ij}^A \tag{11}$$

Foreground-to-foreground entropy: $H_{F \to F}(t)$

$$= -\frac{1}{2} \sum_{i=t+1}^{N_G} \sum_{j=t+1}^{N_G} p_{ij}^D \log p_{ij}^D \tag{12}$$

$$H_{\text{local}}(t) = H_{B \to B}(t) + H_{F \to F}(t) \tag{13}$$

The local entropy method seeks to select a threshold that maximizes $H_{\text{local}}(t)$. The components of $H_{\text{local}}(t)$ measure the relative probability of transitions from background-to-background or foreground-to-foreground, so maximizing the local entropy will favor thresholds that group similar gray-scale values together.

Joint entropy method

Quadrant B and C in the transition probability distribution provide information on gray-scale transitions from background-to-foreground and foreground-to-background, respectively. Joint entropy (Pal and Pal, 1989) is defined as

Background-to-foreground entropy: $H_{B \to F}(t)$

$$= -\frac{1}{2} \sum_{i=0}^{t} \sum_{j=t+1}^{N_G} p_{ij}^B \log p_{ij}^B \tag{14}$$

Foreground-to-background entropy: $H_{F \to B}(t)$

$$= -\frac{1}{2} \sum_{i=t+1}^{N_G} \sum_{j=0}^{t} p_{ij}^C \log p_{ij}^C \tag{15}$$

$$H_{\text{joint}}(t) = H_{B \to F}(t) + H_{F \to B}(t) \tag{16}$$

The joint entropy method seeks to select a threshold that maximizes $H_{\text{joint}}(t)$. The components of $H_{\text{joint}}(t)$ measure the probability of transitions between pixels in different segments. Maximizing this value favors thresholds that minimize the likelihood of pixels with similar gray-scale values being grouped together.

Relative entropy method

The relative entropy method attempts to find a threshold that minimizes the mismatch between the transition probability distributions of the thresholded image and those of the original gray-scale image (Chang *et al.*, 1994). The smaller the relative entropy,

the better the match between the gray-scale image and the corresponding thresholded image.

The dimensions of the co-occurrence matrix and the image are $N_G \times N_G$. Because there are only two gray-scales in the binary image, equal probability is applied in each quadrant of the co-occurrence matrix. The following are the definitions for the probability distributions in each quadrant for a binary image.

$$p_{ij}^{\prime(A)} = q_A(t) = \frac{P_A(t)}{(t+1)\times(t+1)} \quad \text{for } 0 \le i \le t,\ 0 \le j \le t \tag{17}$$

$$p_{ij}^{\prime(B)} = q_B(t) = \frac{P_B(t)}{(t+1)\times(N_G - t)} \quad \text{for } 0 \le i \le t,\ t+1 \le j \le N_G \tag{18}$$

$$p_{ij}^{\prime(C)} = q_C(t) = \frac{P_C(t)}{(N_G - t)\times(t+1)} \quad \text{for } t+1 \le i \le N_G, \quad 0 \le j \le t \tag{19}$$

$$p_{ij}^{\prime(D)} = q_D(t) = \frac{P_D(t)}{(N_G - t)\times(N_G - t)} \quad \text{for } t+1 \le i \le N_G,\ t+1 \le j \le N_G \tag{20}$$

The relative entropy based on co-occurrence matrix is defined as

$$H_{\text{relative}}(t) = \sum_{i=0}^{N_G}\sum_{j=0}^{N_G} p_{ij} \log \frac{p_{ij}}{p'_{ij}} = \sum_{i=0}^{N_G}\sum_{j=0}^{N_G} p_{ij} \log p_{ij} - \sum_{i=0}^{N_G}\sum_{j=0}^{N_G} p_{ij} \log p'_{ij} \tag{21}$$

The goal of minimizing $H_{\text{relative}}(t)$ is equivalent to maximizing the second term of the above equation

$$\begin{aligned}\sum_{i=0}^{N_G}\sum_{j=0}^{N_G} p_{ij} \log p'_{ij} &= \sum_{A} p_{ij} \log q_A(t) + \sum_{B} p_{ij} \log q_B(t) + \sum_{C} p_{ij} \log q_C(t) + \sum_{D} p_{ij} \log q_D(t) \\ &= p_A(t) \log q_A(t) + p_B(t) \log q_B(t) + p_C(t) \log q_C(t) + p_D(t) \log q_D(t)\end{aligned} \tag{22}$$

The optimal threshold by the relative entropy method is found by maximizing this equation with respect to t.

HISTOGRAM BASED METHODS

Renyi's entropy

Renyi's entropy (Sahoo *et al.*, 1997) is based on the histogram of an image rather than on the co-occurrence matrix. It is a combination of the maximum entropy sum method and the entropic correlation method. In Renyi's entropy, the probability distributions of the foreground and background classes, Λ_1 and Λ_2, are defined for the threshold t as

$$\Lambda_1 : \frac{p_0}{p(\Lambda_1)}, \quad \frac{p_1}{p(\Lambda_1)}, \ldots, \frac{p_t}{p(\Lambda_1)} \tag{23}$$

and

$$\Lambda_2 : \frac{p_{t+1}}{p(\Lambda_2)}, \quad \frac{p_{t+2}}{p(\Lambda_2)}, \ldots, \frac{p_{N_G}}{p(\Lambda_2)} \tag{24}$$

where

$$p(\Lambda_1) = \sum_{i=0}^{t} p_i, \quad p(\Lambda_2) = \sum_{i=t+1}^{N_G} p_i, \quad \text{and } p(\Lambda_1) + p(\Lambda_2) = 1 \tag{25}$$

Renyi's entropy of order α for an image is given by

$$H(t)^{\alpha} = \frac{1}{1-\alpha} \ln \sum_{k=0}^{t} (p_k)^{\alpha} \tag{26}$$

where $\alpha \neq 1$ and is a real, positive number.

Renyi's entropies, for the background and foreground distributions, are given by

$$H(t)^{\alpha}_{\Lambda_1} = \frac{1}{1-\alpha} \ln \sum_{i=0}^{t} \left(\frac{p_i}{p(\Lambda_1)}\right)^{\alpha} \tag{27}$$

and

$$H(t)^{\alpha}_{\Lambda_2} = \frac{1}{1-\alpha} \ln \sum_{i=t+1}^{N_G} \left(\frac{p_i}{p(\Lambda_2)}\right)^{\alpha} \tag{28}$$

Let $t^*(\alpha)$ be the threshold level that maximizes the sum of the Renyi entropies

$$t^*(\alpha) = \max\left\{H(t)^{\alpha}_{\Lambda_1} + H(t)^{\alpha}_{\Lambda_2}\right\} \tag{29}$$

where $t^*(\alpha)$ is a function of α and it is calculated at three different values according to Sahoo *et al.* (1997).

$$t_1^* = t^*(0.5)$$

$$t_2^* = t^*(1)$$

$$t_3^* = t^*(2)$$

The optimal threshold t_c^* is a function of t_1^*, t_2^*, and t_3^*. It is defined below

$$t_c^* = t_1^*\left[p(t_1^*) + \frac{1}{4}\omega\beta_1\right] + \frac{1}{4}t_2^*\omega\beta_2 + t_3^*\left[1 - p(t_3^*) + \frac{1}{4}\omega\beta_3\right] \tag{30}$$

where

$$p(t_j^*) = \sum_{i=0}^{t} p_i, \quad \text{and } \omega = p(t_3^*) - p(t_1^*) \tag{31}$$

and

$$(\beta_1, \beta_2, \beta_3) = \begin{cases} (1,2,1) & \text{if } |t_1^* - t_2^*| \le 5 \text{ and } |t_2^* - t_3^*| \le 5 \\ (1,2,1) & \text{if } |t_1^* - t_2^*| > 5 \text{ and } |t_2^* - t_3^*| > 5 \\ (0,1,3) & \text{if } |t_1^* - t_2^*| \le 5 \text{ and } |t_2^* - t_3^*| > 5 \\ (3,1,0) & \text{if } |t_1^* - t_2^*| > 5 \text{ and } |t_2^* - t_3^*| \le 5 \end{cases} \tag{32}$$

The optimal threshold t_c^* is bounded by

$$\min(t_1^*, t_2^*, t_3^*) \le t_c^* \le \max(t_1^*, t_2^*, t_3^*)$$

and may be chosen arbitrarily, or on the basis of some additional function which produces the desired results. For example, select one of the bounds, or a linear function of the bounds. In our tests we used the calculated t_c^* value from Eq. (30).

Iterative selection method

The iterative selection method is also based on the histogram of an image but is simpler than Renyi's entropy. It assumes that there is a mean value for both background and foreground pixels (Ridler and Calvard, 1978) and that an optimal threshold can be achieved by averaging these two values. The algorithm for Iterative Selection is

Assume that the range of gray-scale level ranges is [0 .. N_G].
Let $t_c^* = N_G$ and moving_index = 0;
while(moving_index < = t_c^* and moving_index < = N_G)
{
  background_sum= $\sum_{i=0}^{\text{moving_index}} i \times \text{histogram}[i]$
  histogram_sum_background= $\sum_{i=0}^{\text{moving_index}}$ histogram[i]
  foreground_sum= $\sum_{i=\text{moving_index}+1}^{N_G} i \times \text{histogram}[i]$
  histogram_sum_foreground = $\sum_{i=\text{moving_index}+1}^{N_G}$ histogram[i]
  t_c^* =(background_sum ÷ histogram_ sum_ background + foreground_sum + foreground_sum ÷ histogram_sum_foreground) ÷ 2
  increase moving_index by 1
}

The optimal threshold, t_c^*, will be the final threshold calculated. It will be the threshold that provides the maximum degree of separation between the average gray-scale level of the foreground and background segments.

EXPERIMENTAL METHODS

In our previous studies we showed that biofilms grown at different flow velocities and substrate levels had different structures (Beyenal and Lewandowski, 2000). Therefore, biofilms of varying structure were obtained for imaging by growing them at different flow velocities and glucose concentrations. To investigate the effects of microscopy type on the thresholding technique, we used images acquired from transmission and reflected light microscopy, and confocal scanning laser microscopy.

Biofilm growth

The biofilms were grown in an open channel flow reactor, made of polycarbonate (3.5 cm deep, 2.5 cm wide, and 85 cm long with a total working volume of 400 mL). A schematic view of the experimental set-up was presented by Xia *et al.* (1998). The reactor was placed on a Nikon Diaphot 200 inverted microscope for further imaging of the biofilms. The lid of the reactor was sealed with silicon rubber to prevent contamination and the entire system was sterilized with 70% alcohol before each experiment. The reactor and components were rinsed with sterile water (autoclaved) until all the alcohol was removed. The reactor components (tubings, connectors, air filter, etc.) and growth medium were autoclaved at 121°C prior to experiments.

The broth solution was made of KH_2PO_4 (0.35 g/L), Na_2HPO_4 (1.825 g/L), $(NH_4)_2SO_4$ (0.1 g/L), $MgSO_4 \times 7H_2O$ (0.01 g/L), yeast extract (0.01 g/L) and glucose (different concentrations for each experiment). The glucose and yeast extract were autoclaved separately and then added to the growth medium.

The following procedure was used to inoculate the reactors. One mL portions of frozen stock of *Pseudomonas aeruginosa* (ATCC 700829), *Pseudomonas fluorescens* (ATCC 700830), and *Klebsiella pneumoniae* (ATCC 700831) were inoculated to separate flasks (50 mL working volume) including growth medium (0.2 g/L yeast extract, 0.5 g/L glucose and the nutrientes given above). The flasks were then shaken for 24 h and 15 mL of each culture from the flasks were inoculated to the flat plate flow cell. The reactor was operated batch-mode (recycling) for 12 h and than switched to the continuous feed mode. To assure glucose concentration was kept almost constant in the reactors, we measured glucose concentration using procedure 510 by Sigma® Diagnostics.

Peristaltic pumps (Cole-Parmer, Chicago, IL) were used to maintain flow and recycle rates. The recycle ratio (the ratio of volumetric flow rates of recycle and feed) was greater than 30 and the hydraulic retention time was less than 10 min to prevent growth of microorganisms in suspension and to keep the substrate concentrations constant along the reactor. The difference between the influent and effluent glucose concentrations was consistently less than 10% of that in the influent. The recycled solution was continuously aerated in a mixing chamber before entering the reactor. The dissolved oxygen concentration was monitored using a dissolved oxygen electrode and was kept near saturation value.

The biofilms studied by the transmission microscopy were grown at 2.5 cm/s flow velocity and 40 mg/L glucose concentration. Images of the biofilms were captured every day at different locations using 4 × objective. For analysis we used images taken at days 3, 6 and 9. The biofilms studied by the CSLM were grown at 2.5 cm/s flow velocity and 40 mg/L glucose concentration. Analysis was based on image taken at days 5 and 7.

Biofilms studied by the reflected light microscopy were grown at 3.2, 10 and 25 cm/s flow velocity and 50 and 150 mg/L glucose concentrations. Analysis was based on images taken at days 3, 6 and 7.

Biofilm imaging

Images acquired by light microscopy. The flat plate reactor was placed on the Nikon Diaphot 200 inverted microscope. Two different types of illumination were used to obtain two different types of biofilm images; (1) light directed from the bottom of the biofilm (reflected light microscopy images) and (2) light directed from the top of the biofilm (transmission microscopy images). The biofilm was viewed from the bottom using 4 or $10\times$ objectives. Images were captured by a computer integrated with a COHU® camera (Closed Circuit, CA; model no: 2222-1040/0000) and Flashpoint® frame grabber (Integral technologies Inc., Indianapolis, Indiana). The images were taken in TIFF format and consisted of 640×480 pixels using Image Pro® (Media Cybernetics, Maryland).

Images acquired by confocal scanning laser microscopy (CSLM). A Bio-Rad, model MRC600, confocal scanning laser microscope integrated with an Olympus, model BH2, and light microscope was used for biofilm imaging. Detailed procedures of the imaging are given by Xia *et al*. (1998). We used 1% v/v, acridine orange (from Sigma) to stain the biofilm. After staining the biofilm, we acquired confocal images at different locations and distances from the bottom. Ten CSLM images at different depth were captured to test automatic thresholding methods.

Manual and automatic thresholding of the images

Twenty 8-bit gray-scale (256 gray-scale levels) light microscopy biofilm images were used to test the manual and automatic thresholding methods. The image reference set was collected from different experiments and covered a broad spectrum of biofilms of different age, grown at different flow velocities, and at different nutrient concentrations. These provided different structured biofilm images to be analyzed.

Manual thresholding was performed by 10 biofilm researchers from the Center for Biofilm Engineering at Montana State University (Bozeman). Their results were averaged and the standard deviation was calculated to assess the variability of operator choice.

Automatic thresholding was performed using the Image Structure Analyzer (ISA), a software package developed by the Structure-Function group, Center for Biofilm Engineering at Montana State University for quantifying biofilm heterogeneity (Yang *et al*., 2000). It was developed for the UNIX/Motif environment in C++ with all calculations done in double precision arithmetic. All tested thresholding methods were integrated into this package and tested against the reference image set.

Comparison of the automatic thresholding methods

To evaluate the differences between the threshold level selected by automatic thresholding methods and those selected manually, the mean of the sum squares of relative residuals (MSSRR) was calculated as

$$\mathrm{MSSRR}=\frac{1}{N_m}\sum_{i=1}^{N_m}\left(\frac{\mathrm{va}_i-\mathrm{v}_i}{\mathrm{va}_i}\right)^2 \tag{33}$$

where v_i is the threshold of an automatic thresholding method for image i, va_i is the standard (the average of the manual) threshold for the corresponding image, and N_m is the total number of images. When N_m is sufficiently large, the root mean square error gives an average deviation of the threshold level for the method. A value of $N_m=20$ was found to be sufficient to calculate a statistically meaningful MSSRR value using the relative least square method. Sáez and Rittman (1992) demonstrated that when the magnitude of dependent variable (in our case the treshold value) varies widely, using relative residuals method is superior to using absolute residuals. We use the regression coefficient, R^2, to describe the correlation between the manual and automatic thresholds.

RESULTS

An example of automatic thresholding

Figure 2(A) shows a gray-scale biofilm image acquired using light microscopy (the biofilm was grown at 2.5 cm/s flow velocity, at 40 mg/L glucose concentration and was five days old). The darker areas are biomass and the bright areas are interstitial space. Five threshold methods were used and the results are shown in Fig. 2(B)–(F). The thresholded images are presented as black (biomass) and white (interstitial space). The iterative selection method showed the best agreement with the manual thresholding shown in Fig. 2(G).

Gray-scale biofilm images

Figure 3 compares the threshold levels selected automatically and manually for 20 biofilm images. Results obtained using the iterative selection method (Fig. 3(A)), local entropy method (Fig. 3(B)), joint entropy method (Fig. 3(C)), relative entropy method (Fig. 3(D)), and Renyi's entropy (Fig. 3(E)) are plotted against the average value obtained from manual thresholding. Error bars show the respective standard deviations from the average. The figure shows that local entropy, joint entropy, relative entropy, and Renyi's entropy methods yield results that are poorly correlated with the results of manual thresholding. A good correlation between the results of automatic and manual thresholding was found only for the iterative selection method ($R^2=0.96$). Table 1 reports the MSSRR and correlation coefficients between the automatic and manual thresholding method results.

Confocal images

Images of biofilms taken with a CSLM have a minimal range of gray levels. As a result, the images are nearly binary without thresholding (Lawrence *et al*., 1991). However, the threshold selected by iterative selection again produced the best correlation with manual thresholding (results not shown).

DISCUSSION

Accurate measurements of image features using image analysis require an accurate, reliable method of image thresholding. Ideally, an automated method should be used to avoid individual preferences of the operators. We have tested five threshold algorithms using images of biofilms of different physical structure and age. The images were acquired using different microscopy techniques. To determine the accuracy of the selected thresholding methods, we

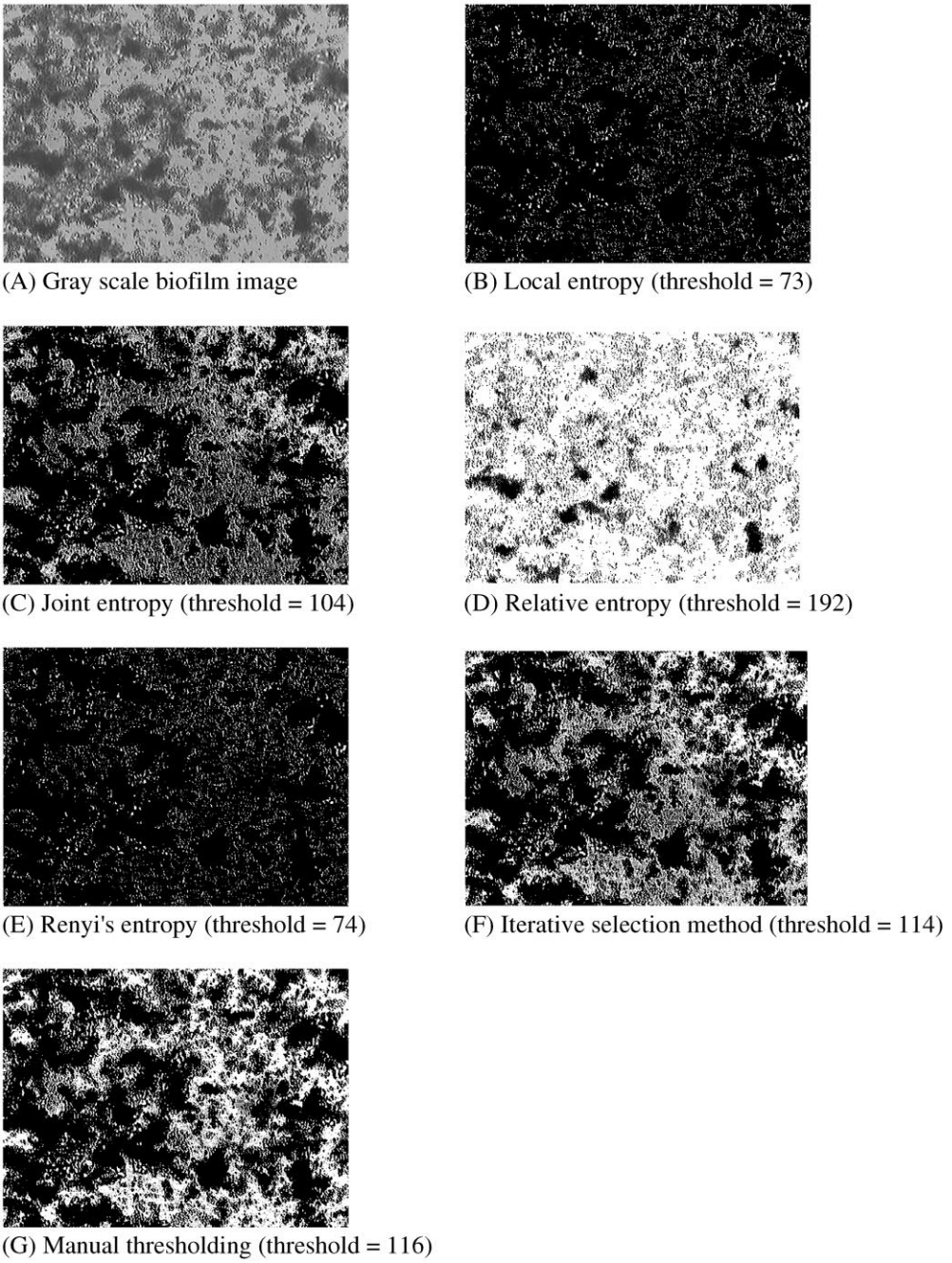


Fig. 2. A gray-scale biofilm image (A), thresholded by different methods: local entropy (B); joint entropy (C); relative entropy (D); Renyi's entropy (E); iterative selection (F); the average value of manual thresholding (G).

assumed that the correct value of the threshold is that selected manually by a team of expert biofilm researchers. The accuracy of the tested methods was then estimated as the deviation between the threshold estimated using the respective algorithms and the correct threshold. The thresholds determined by the iterative selection method were most accurate.

Attempts to select the best thresholding technique often define the correct value of the measured quantity and then evaluate the thresholding procedures based on how closely they come to the true value. In our study we used the average threshold from those estimated by several biofilm researchers. Sieracki *et al.* (1989) tested nine threshold techniques all based on the edge detection and used fluorescent beads of known size to determine the correct treshold. Dubuission *et al.* (1994) also used edge detection techniques for thresholding of bacterial

culture images but did not define the correct value of the measured parameters, and therefore, could not estimate the accuracy of the applied procedures.

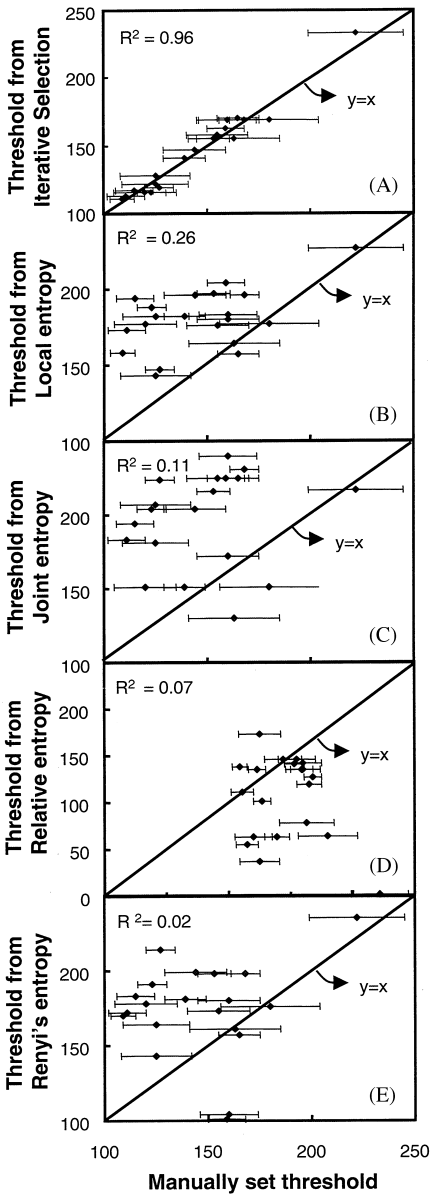


Fig. 3. Results of automatic thresholding compared with those set manually: (A) iterative selection; (B) local entropy; (C) joint entropy; (D) relative entropy; (E) Renyi's entropy. The results of manual thresholding are shown as an average along with standard deviations from the average (error bars). For an agreement between automatic and manual thresholding we expect that the data fall on the diagonal line ($y = x$).

There are distinct differences between our work and similar works described in literature. Edge detecting techniques described in the above papers were performed on systems where the edges were easily detectable: fluorescent beads or microbial cells. The known size of the fluorescent beads could also serve as a convenient reference to estimate the accuracy of applied image analysis procedures. In our case, however, cell clusters in biofilms did not have distinctly defined edges and their exact size was unknown. Consequently, the reference used to evaluate the accuracy of thresholding procedures had to be established based on manual thresholding aided by appropriate statistical procedures.

The ultimate goal of our work is to develop a software package that automatically evaluates several parameters from biofilm images (see our recent paper: Yang *et al*., 2000). Although some quantitative parameters characterizing image structure can be extracted from microscope images using commercial software packages for image analysis, most of the software uses manual thresholding.

CONCLUSIONS

Comparing the results of manual thresholding with those of automatic thresholding indicates that iterative selection is superior to the other automatic methods tested. It is also the only automatic method that is sufficiently precise to be considered as a potential replacement for manual thresholding. The results were similar for images collected by light or confocal microscopy. These results are of interest to anyone using thresholding in biofilm image analysis, and may well be extensible to other types of image analysis, such as medical imaging and automated manufacturing.

Acknowledgements—The research was supported by the cooperative agreement EED-8907039 between the National Science Foundation and Montana State University. Special thanks to Professor Martin Hamilton for his help with the statistical analysis. The authors also thank researchers from the Center for Biofilm Engineering at Montana State University (Bozeman) who helped with the manual thresholding.

REFERENCES

Beyenal H. and Lewandowski Z. (2000) Combined effect of substrate concentration and flow velocity on effective diffusivity in biofilms. *Water Res*. **34**, 528–538.

Beyenal H., Tanyolac A. and Lewandowski Z. (1998) Measurement of local effective diffusivity in heterogeneous biofilms. *Water Sci. Technol*. **38**, 171–173.

Table 1. MSSRR and correlation coefficients between the results of automatic and manual thresholding

	Local entropy	Joint entropy	Relative entropy	Renyi's entropy	Iterative selection
MSSRR	0.111	0.204	0.172	0.137	0.001
R^2	0.26	0.11	0.07	0.02	0.96

Bishop P. L. and Rittmann B. E. (1995) Modelling heterogeneity in biofilms: report of the discussion session. *Water Sci. Technol.* **32**, 263–265.

Brink A. D. (1992) Thresholding of digital images using two-dimensional entropies. *Pattern recognition* **25**, 803–808.

Chang C. I., Chen K., Wang J. and Althouse M. L. G. (1994) A relative entropy-based approach to image thresholding. *Pattern recognition* **27**, 1275–1289.

Chang F. J., Yen J. C. and Chang S. (1995) A new criterion for automatic multilevel thresholding. *IEEE Trans. Image Process* **4**, 370–378.

Costerton J. W., Lewandowski Z., DeBeer D., Caldwell D., Korber D. and James G. (1994) Minireview: Biofilms, the customized microniche. *J. Bacteriol.* **176**, 2137–2142.

Dubuission M-P., Jain A. K. and Jain M. K. (1994) Segmentation and classification of bacterial culture images. *J. Microbiol. Meth.* **19**, 279–295.

Haralick R. M., Shanugam K. and Dinstein I. (1973) Textural features for image classification. *IEEE Trans. Systems Man Cybernet.* **SMC-3**, 610–621.

Keevil C. W. and Walker J. T. (1992) A Normarski DIC microscopy and image analysis of biofilms. *Binary* **4**, 93–95.

Lawrence J. R., Korber D. R., Hoyle B. D., Costerton J. W. and Caldwell D. E. (1991) Optical sectioning of microbial biofilms. *J. Bacteriol.* **173**, 6558–6567.

Lee S. U. and Chung S. Y. (1990) A comparative performance study of several global thresholding techniques for segmentation. *Comput. Vision Graphics Image Process.* **52**, 171–190.

Lewandowski Z., Webb D., Hamilton M. and Harkin G. (1999) Quantifying biofilm structure. *Water Sci. Technol.* **39**, 71–76.

Massol-Deya A. A., Whallon J., Hickey R. F. and Tiedje J. M. (1994) Channel structures in aerobic biofilms of fixed-film reactors treating contaminated groundwater. *Appl. Environ. Microbiol.* **61**, 769–777.

Noguera D. R., Okabe S. and Picireanu C. (1999) Biofilm modeling: Present status and future directions. *Water Sci. Technol.* **39**, 273–278.

Pal N. R. and Pal S. K. (1989) Entropy thresholding. *Signal process.* **16**, 97–108.

Ridler T. W. and Calvard S. (1978) Picture thresholding using an iterative selection method. *IEEE Trans. Systems Man Cybernet.* **SMC-8**, 630–632.

Rittmann B. R., Pettis M., Reeves W. H. and Stahl D. (1999) How biofilm clusters affect substrate flux and ecological selection. *Water Sci. Technol.* **39**, 99–105.

Sáez P. B. and Rittmann B. E. (1992) Model parameter estimation using least squares. *Water Res.* **26**, 489–796.

Sahoo P., Wilkins C. and Yeager J. (1997) Threshold selection using Renyi's entropy. *Pattern Recognition* **30**, 71–84.

Sieracki M. E., Reichenbach S. E. and Webb K. L. (1989) Evaluation of automated threshold selection methods for accurately sizing microscopic flourescent cells by image analysis. *Appl. Environ. Microbiol.* **55**, 2762–2772.

Wilkinson M. H. F. (1998) Automated and manual segmentation techniques in image analysis of microbes. In *Imaging, Morphometry, Fluorometry and Motility Techniques and Applications*, eds M. H. F. Wilkinson and F. Schut. Wiley, New York.

Wolfaardt G. M., Lawrence J. R., Robarts R. D. and Caldwell D. E. (1994) A multicellular organization in a degradative biofilm community. *Appl. Environ. Microbiol.* **60**, 434–446.

Xia F., Beyenal H. and Lewandowski Z. (1998) An electrochemical technique to measure local flow velocity in biofilms. *Water Sci. Technol.* **32**, 3631–3636.

Yang S. and Lewandowski Z. (1995) Measurement of local mass transfer coefficient in biofilms. *Biotechnol. Bioeng.* **48**, 737–744.

Yang X., Beyenal H., Harkin G. and Lewandowski Z. (2000) Quantifying biofilm structure using image analysis. *J. Microbiol. Meth.* **39**, 109–119.