Available online at www.sciencedirect.com
SCIENCE DIRECT
Neural Networks 18 (2005) 850–860
Neural Networks
www.elsevier.com/locate/neunet

2005 Special Issue

Refining competition in the self-organising tree map for unsupervised biofilm image segmentation

Matthew Kyan[a,*], Ling Guan[b], Steven Liss[c]

[a]School of Electrical and Information Systems Engineering, University of Sydney, NSW 2006, Australia
[b]Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON, Canada M5B2K3
[c]Department of Chemistry and Biology, Ryerson University Toronto, ON, Canada M5B2K3

Abstract

The Self Organising Tree Map (SOTM) neural network is investigated as a means of segmenting micro-organisms from confocal microscope image data. Features describing pixel and regional intensities, phase congruency and spatial proximity are explored in terms of their impact on the segmentation of bacteria and other micro-organisms. The significance of individual features is investigated, and it is proposed that, within the context of micro-biological image segmentation, better object delineation can be achieved if certain features are more dominant in the initial stages of learning. In this way, other features are allowed to become more/less significant as learning progresses: as more knowledge is acquired about the data being segmented. We argue that the efficiency and flexibility of the SOTM in adapting to, and preserving the topology of input space, makes it an appropriate candidate for implementing this idea.

We propose a refinement to the competitive search strategy that allows for a more appropriate fusion of signal and proximal features, thereby promoting a segmentation that is more sensitive to the regional associations of different microbial matter. A refined stop criterion is also suggested such that the dynamically generated number of classes becomes more data dependant. Preliminary experiments are presented and it is found that favouring intensity characteristics in the early phases of learning, whilst relaxing proximity constraints in later phases of learning, offers a general mechanism through which we can improve the segmentation of microbial constituents.[1]

1. Introduction

Microscopy has long been considered an essential investigative tool in the study of complex micro-biological environments, such as those encountered in modern Biofilm research. The term *Biofilm* is used to describe a collection of micro-organisms anchored to some form of substrata (Palmer & Sternberg, 1999), which may exist in a variety of different situations—from the dental plaque found on teeth, to microbial *flocs* (highly complex aggregations of microbial cells or bacteria, bioorganic and inorganic material) suspended in aqueous environments (Liss, 2002). In such environments, physical structure plays a crucial role in the many key processes that occur.

Knowledge of the structural factors affecting such processes becomes important in formulating solutions to many pertinent environmental problems. In the performance of waste water treatment, for instance, a process known as *activated sludge* is utilised to promote the stabilised formation of flocs. This process then aids in the separation of suspended microbial contaminants from treated effluent. A common problem occurs in such systems when flocs form with poor settling properties, thereby degrading the overall treatment process (Liss, 2002). To gain insight into the structural properties and microbial behaviour in this and many other environmental applications, microscopy remains the major tool through which non-invasive study may be conducted in situ at the cellular level.

Major technological advances in microscopy such as the introduction and development of the confocal laser scanning microscope (CLSM), can now overcome many of the normal problems associated with using standard instrumentation to capture images of these, typically dense,

* Corresponding author.
E-mail addresses: mkyan@ee.ryerson.ca (M. Kyan), lguan@ee.ryerson.ca (L. Guan), sliss@ryerson.ca (S. Liss).

[1] An abbreviated version of some portions of this article appeared in Kyan, Guan, and Liss (2005) [Kyan, M., Guan, L., & Liss, S. (2005). *Dynamic feature fusion in the self organising Tree map—applied to the segmentation of biofilm images. Proceedings of the International Joint Conference on Neural Networks. July 31–August 4, 2005, Montreal, Canada.*], published under the IEEE copyright.

0893-6080/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2005.06.032

heterogeneous environments. The ability for CLSM to probe into the depth of a specimen and record a stack of 2D spatially registered image slices allows for reconstruction of the 3D structural information microbiologists and engineers need to form better models of the biofilms or flocs under investigation. At this point, biofilm structure is understood more in a qualitative rather than a quantitative sense (Beyenal, Lewandowksi, & Harkin, 2004). As such, the linking of parameters characterising biofilm performance, with structural features remains a hot topic of research.

In this work, we identify that much of the current approaches to quantifying biofilm structure hinge on consideration of global characteristics. Such measures break down when the image data is extremely heterogenous due to information lost through crude segmentation that is not representative of the internal components of the data. This factor is particularly evident in images of microbial floc. We thus present a neural network approach to the segmentation of the internal constituents of biofilm/floc images—incorporating and fusing a set of image features that together, cover a range of key aspects of the human visual system (HVS). Finally, we introduce a restriction on the scope of resolution considered in the feature space in order that the dynamic generation of classes may be bounded by the data. We then propose the combination of this with a mechanism for the dynamic adjustment of input feature significance. Together, they help to guide the SOTM as it searches through feature space for a better perceptual representation of the input data.

2. Theoretical considerations

At present, the majority of available techniques attempt to quantify biofilm structure from microscope images by segmenting the image into foreground and background pixels. Most frequently reported techniques in use within microbiological applications are based on more traditional image processing approaches that utilise boundary and edge information. Such techniques however, are typically tested only in situations where microbial constituents are clearly defined—for example, in the imaging of microbes specifically labelled with fluorescent particles (Wilkinson, 1998). These approaches often become ineffective when applied to biofilm, in which objects are not clearly defined, often existing within a dense, spatially heterogenous mixture of bacterial species, extracellular polymeric substances (EPS), and other organic material. This heterogeneity is further magnified in the case of floc studies.

Automatic techniques for thresholding have been proposed for use with biofilm. A series of entropy and histogram based approaches are reviewed in the literature (Yang, Beyenal, Harkin, & Lewandowski, 2001), for their applicability to biofilm. Some of the more well known methods include iterative selection Ridler and Calvard (1978), and Otsu's method (Otsu, 1979)—as utilised by the popular biofilm characterisation software ISA5 (developed by the Center for Biofilm Engineering at Montana State University). There is some debate within the biofilm community however, as to what constitutes a more appropriate and robust method of biofilm segmentation (Baveye, 2002), arguing in favour of alternative methods such as the minimum-error algorithm (Kittler & Illingworth, 1986), or the ICM: Iterated Conditional Mode (Besag, 1986) algorithm.

At this point, the majority of approaches explored thus far treat biofilm images as essentially bimodal in nature. Whilst this might be a reasonable assumption for some experiments where the biofilm is relatively homogeneous, it is rarely the case in the study of floc. Presently, multimodal or multilevel considerations have received little attention in the field.

It should be acknowledged that quite often, epi-fluorescent modes of the microscope are used to image the distribution of molecular markers introduced to a specimen. In such applications, the markers bond with particular molecules present in structures of interest so that they can be isolated from their surroundings. However, often these molecules exist in many different materials present, and thus, are generally present in varying concentrations, reflected as a range of graylevels in the image.

In such instances, segmentation via crude thresholding of foreground from background will result in the loss of much, potentially significant information. As a result, often researchers resort to the arduous task of manual or semi-automated segmentation—where operator discretion plays a key role in impacting any subsequent characterisation measurement taken.

In addition, popular quantitative approaches generally measure morphological and textural parameters (Beyenal et al., 2004; Tolle, 2003). Morphological measures discussed in the literature are dependant almost exclusively on an initial thresholding. Any form of particle property measured (porosity, size, etc) must then assume that the particles existing in the 'foreground' are of the same type—when in fact they may be a conglomerate of overlapping types, with different intensity profiles or shapes. In the context of a thresholded 'biomass', separation of particles is difficult without consideration of the inherent heterogeneity in the data (Fig. 1). Fractal dimension is a common property discussed in the literature. It is generally taken over the entire field of view (FOV), or a set of localised windows regularly spaced over the FOV. Again, in any one FOV taken, often the mixture of material present is not homogeneous, thus the measure is contaminated or diluted by the underlying heterogeneity.

Current approaches taken to quantify textural parameters from biofilm images also suffer, as they tend to involve the capture of some global measure over the entire FOV. This effectively smooths over any true heterogeneity existing in the data. If the material present in the biofilm or floc images could be better separated—such measurements could be

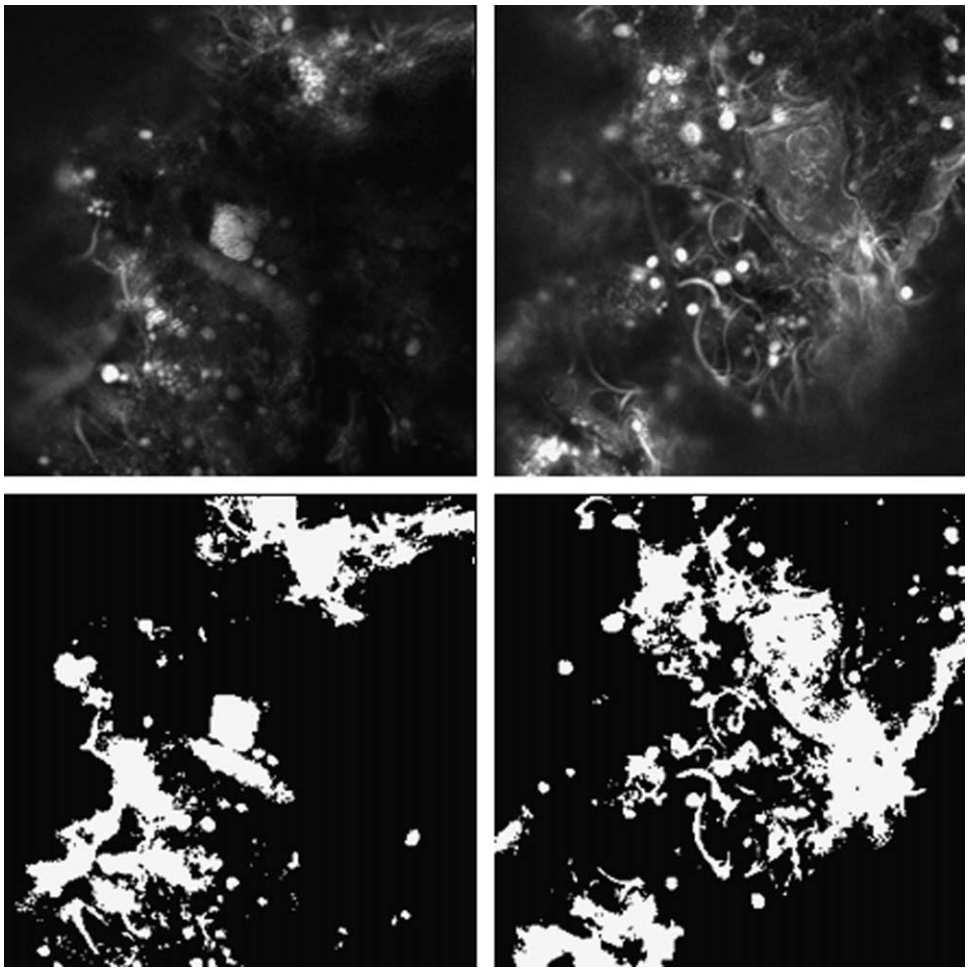

Fig. 1. Two slices of FITC-stained biofilm image data (above), and associated Otsu thresholded binary representations (below). Low contrast 'hazes' indicate of trace distributions of fluorescent molecular markers -often not noise, yet are truncated through thresholding, whilst clusters of cells are blended together making subsequent particle analysis difficult.

localised to the regions in which a specific type of material exists. This would result in a more accurate quantification of the nature and distribution of individual materials existing within a biofilm sample.

As such, a fundamental step toward meaningful characterisation lies in investigating better forms of segmentation. The goal being firstly, to achieve more robust and repeatable isolation of biofilm/floc structure from a background of noise imparted through the imaging optics; secondly, to isolate the internal constituents of floc, such that parameters measured may more appropriately characterise the distribution and nature of materials present.

3. Segmentation with the SOTM neural network

In this current work, a tree structured neural network known as the Self Organising Tree Map (SOTM) is investigated and developed toward the goal of achieving better segmentation of biofilm constituents. The SOTM partitions a feature space description of input data, by locating clusters (represented by neurons) of high density within this feature space. Like its counterpart, the self organising map (SOM) (Kohonen, 1982), competitive learning is used to locate clusters such that the final representation maintains the topology of the feature space, yet does so in a more flexible and efficient manner by dynamically generating a model of this topology as it parses the input space. This results in a representation that tends not to suffer (as SOM does), from neurons being trapped in regions of low density (Kong, 1998; Randall, Guan, Zhang, & Li, 1999).

3.1. The SOTM algorithm

The process for SOTM learning is outlined below:

(a) Initialise the root neuron with a randomly selected training vector from feature space $\mathbf{X}$.
(b) Randomly select a new feature vector $\mathbf{x}$, and compute the Euclidean distance, d_j, to all currently existing neurons $\mathbf{w}_j$ ($j=1,\ldots,J$), where J is the total number of neurons currently allocated

(c) Select the winning neuron $\mathbf{w}_k$, with minimum d_j: $d_j \cdot (\mathbf{x}, \mathbf{w}_k) = \min_j d_j \cdot (\mathbf{x}, \mathbf{w}_j)$
(d) If $d_j \cdot (\mathbf{x}, \mathbf{w}_k) \leq H(t)$, then update the weight vector of the kth neuron using the reinforced learning rule: $\mathbf{w}_k(t+1) = \mathbf{w}_k(t) + \alpha(t)[\mathbf{x} - \mathbf{w}_j]$; where $\alpha(t)$ is the learning rate and $H(t)$ is a hierarchy control function. Both functions decay with time: $\alpha(t)$ lies on [0,1], whilst $H(t)$ decays exponentially, controlling the levels of the tree.
(e) Alternatively, if $d_j \cdot (\mathbf{x}, \mathbf{w}_k) > H(t)$, spawn a new neuron from $\mathbf{w}_k$ at the position $\mathbf{x}$ and reset $\alpha(t)$.
(f) Continue from step (b) until all neurons are allocated and/or there is no significant change in the SOTM.

As indicated above, the network is initialised with the random selection of a vector from input space (stimulus): typically each pixel will be represented by a single input vector. The vector may be as simple as a single graylevel intensity value, or it may be a more complex multi-dimensional vector with each dimension representing a different pixel feature (such as some form of local statistic or otherwise).

Learning progresses through the random selection of new stimuli from the feature space. With each, the currently existing neurons in the network compete to see which is the closest to the stimuli in a Euclidean sense. We can think of this as stimulation of the most similar representation in the memory of the network.

Once a neuron is stimulated, a decision is made as to whether the currently chosen neuron is significantly different from the input or not. This typically occurs in the earlier stages of network growth when there are fewer neurons that exist to compete. The hierarchy control function (described in Section 3.2) is responsible for making this decision. If a neuron is found to exist that is significantly similar to the stimulus, it 'learns' some information from that stimulus. This is achieved by adjusting the neuron (memory) vector toward the current stimulus by a certain proportion $\alpha(t)$ of the Euclidean distance between them.

If there is no significant similarity, then the network figures that it needs to allocate a new neuron (cluster) to the network topology. This neuron then becomes a child of the neuron it was found to be closest to. Over time, as new neurons are allocated to represent the input, more competition for each new stimulus occurs.

The learning rate $\alpha(t)$ is allowed to decay over time as more neurons are allocated—thus ensuring convergence in that the neurons will eventually settle and 'tune' themselves to represent the centres of dense clusters occurring in the feature space.

The hierarchical function also decays, allowing for neurons to be allocated as leaf nodes of their closest neurons from the previous generation (previous state of the network). Thus the SOTM forms a flexible tree structure that spreads and twists across the feature space, converging to what is essentially, a topologically aware k-means representation.

3.2. SOTM partitioning of feature space

As mentioned in Section 3.1, the generation of new neurons is controlled in essence by an exponentially decaying hierarchy control function $H(t)$. It acts more or less as a radius (or more correctly, an ellipsoid) of significance over the feature space: emanating from the neuron found to be closest to the current stimulus. This is depicted in the two and three-dimensional feature space examples of Fig. 2.

Fig. 2. (a) 2D (left) and (b) 3D (right) feature spaces. **x**'s represent initial stimuli falling within the radius/ellipsoid of significant similarity.

If the current stimulus is within the range defined by this radius (within the ellipsoid boundary), then it is considered significantly similar, and that neuron learns some information from the stimulus. If the stimulus occurs outside the boundary of significant similarity, then the network will spawn a new neuron, attaching it to its closest 'parent' neuron.

The partitioning that ensues can be visualised in Fig. 2(a) . With the root neuron allocated, stimuli falling within a large radius of significance cause the root neuron to perform a random walk about the centre of density of the entire feature space (i.e. its expected value).

When a random stimulus eventually falls outside the radius, a new neuron is spawned and the feature space is instantly partitioned into regions A and B. With this partitioning, the two neurons now compete, and reconverge to their new regional centres of density.

The timing diagram in Fig. 3 shows an example of how the hierarchical function decays, and subsequent neurons are generated. A stepped exponential (Kong, 1998) was found in the current work, to be the more desirable option for $H(t)$. This allows for an adequate search of the input space at a fixed radius, before considering a new range of candidates for neuron generation. Existing neurons are thus given time to readjust to the latest partitioning.

It is noted at this point, for reference, that in Section 4.2, we shall suggest what should happen when we bias

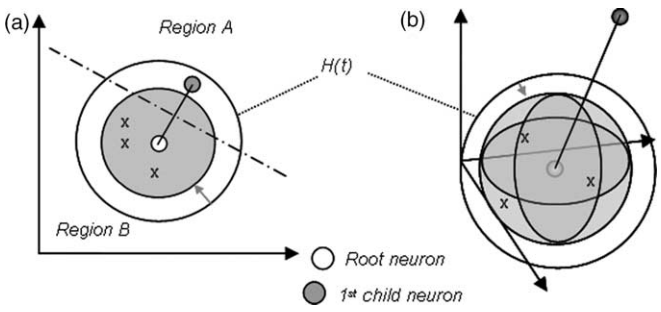

Fig. 2. (a) 2D (left) and (b) 3D (right) feature spaces. **x**'s represent initial stimuli falling within the radius/ellisoid of significant similarity.

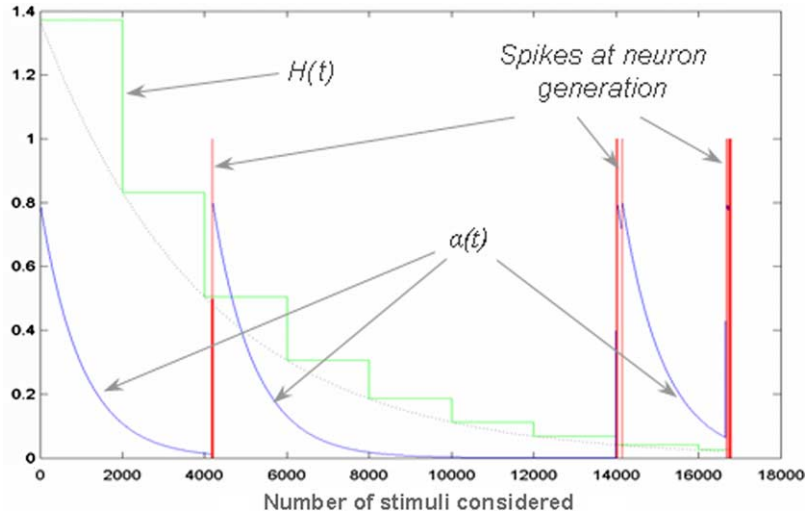

Fig. 3. Timing waveforms governing the evolution of the SOTM.

the Euclidean distance calculation considered in the competitive learning phase, in any one feature dimension. At this point, we might expect that if one particular dimension is weighted down with respect to others, then stimuli distributed along that dimension would be considered to be much closer to the closest neuron than they actually are. As a result, partitioning along other dimensions would occur much more readily.

4. SOTM model refinement

4.1. Automatic class selection

In decaying from a fixed value greater than the maximum possible range across the input feature space, $H(t)$ essentially offers a mechanism by which the resolution of feature space is systematically explored. If $H(t)$ is allowed to decay indefinitely, and there is no limitation on the number of neurons that may be allocated to the network, then it follows that the network will continue to grow indefinitely. Intuitively, this growth should be limited by the smallest resolution of feature points (i.e. the Euclidean sum of the smallest resolutions of all feature sets included in the feature space). Beyond this limit, it makes no sense to generate new cluster centres-yet at this limit, the purpose of clustering to reduce the dimensionality of the data (quantisation), also becomes meaningless, since the network will continue to grow until the number of classes equals the number of feature data points (or greater).

If a meaningful limit can be enforced on what constitutes a reasonable resolution in feature space, then the network would cease to grow once that resolution is achieved and the class centre locations reach stable conditions. Assuming we then have a number of different experimental datasets (images in our case), that are acquired under similar conditions, and defined with similar feature spaces, then the number of classes that are subsequently discovered by the network will be determined by the nature of the data itself.

So, what then constitutes an appropriate minimal resolution in feature space? Well, we know that it must be somewhere between the maximum range of the feature space (Euclidean sum of the range over each individual feature), and the smallest separation between feature points (Euclidean sum of the resolutions of each individual feature) . If the range differs between data sets, then choosing a percentage of this range is not really appropriate. We would prefer a common, fixed resolution that applies to a set of commonly acquired data. Taking a multiple of the minimum resolution in feature space may be a more meaningful choice.

In this way, $H(t)$ may decay from different initial values that are dependant on the natural ranges found in the data, stabilising at a value that is independent of the data (but related to the features chosen), hence common across different data sets. Different centres will be spawned during this process according to how the data are distributed across this range. We thus essentially parse the feature space, allowing the natural distribution of data to trigger the partitioning process.

In light of this, we introduce a new criterion to the stop condition of the network: namely that $H(t)$ must not fall below a certain minimum resolution. In the experiments presented, we limit $H(t)$ to an integer multiple of the minimum possible resolution in feature space.

4.2. Dynamic weighting of feature significance

As a final step, based on the results of applying fixed global weightings to each of the above features for SOTM clustering, we suggest that the network's ability to cluster across an input data space—whilst efficient and optimal in a vector quantisation sense—may not necessarily converge to

locate clusters that represent objects as perceived by the HVS. We propose that in the natural course of learning, certain features may become more (or less) significant. This would have the effect of a non-linear biasing or warping of the final topology of the network away from the vector quantised topology. The resulting topology potentially offers a more object/region based partitioning of the input feature space, and thus, more of an object based segmentation suitable for the delineation of constituent materials existent in biofilm/floc images.

Theoretically, there are quite a number of parameters one could vary in the competitive search, learning, hierarchical function decay, or clustering phases of SOTM segmentation. One might even apply weightings to selectively control all such parameters differently. In this work however, we introduce a dynamic set of weights in the competitive search phase of learning only. We start by applying a set of initial feature weights (β), to the competitive Euclidean distance measure in the SOTM algorithm (Section 3.1) such that step (c) now becomes:

$$d_j : d_j \cdot (x, w_k) = \min_j [\beta \cdot \gamma(t) \cdot (d_j(x, w_j))] \tag{1}$$

where $\gamma(t)$ denotes a set of exponential functions that force individual feature weights to grow or decay on the interval [0 1], or remain constant throughout learning.

More specifically, we utilise a sigmoid function to control the gradual switching of individual weights from one state to another:

$$\gamma(t, c_s, g, T) = \frac{1}{1 + e^{-(q/T)[g(T/2-(t-c_s))]}} \tag{2}$$

where: t is the current cycle number (each presentation of a new input vector to the SOTM represents one cycle); c_s is the delay in cycles before switching should commence; g takes on a value of either $(-1,0,1)$, controlling whether the function decays, remains constant, or grows (respectively) over time; T is the transition period (in cycles) over which the switching occurs; and finally, q controls the shape of the transition profile.

It follows that if individual weights are dynamically increased or decreased gradually (toward the weights of the other features) over the entire learning process, then partitioning would begin by favouring the highest weighted feature(s), shifting toward the other features as they, in turn, become (relatively) more significant.

If feature weights cross over in value, the ratio between them would also switch, thus promoting the shift of partitioning from one feature to another. During this progression, neurons are being allocated in a manner that would favour the highest weighted feature dimension at any one time during learning. This might be a significant dynamic in terms of object segmentation, where neurons to be allocated could favour one dimension until sufficient information is learnt from that dimension, before shifting focus to another dimension.

Such concepts are explored in the following experiments (outlined in Section 6) to draw conclusions as to how such dynamics might result in more effective object segmentation as applied to complex, heterogeneous microbiological image data.

5. Feature selection and fusion

In exploring the effectiveness of the SOTM for segmentation, a set of features is fed into the network, each describing a different aspect of the input data. In this work we have chosen a small set of what the authors believe are features fundamental to a basic description of objects in a graylevel scene, as segmented by the HVS. These features include: Graylevel (GL) of an individual pixel, which describes highly localised intensity variation within an object; Average graylevel of a disk shaped region (GL_{dsk}) surrounding the current pixel, which describes more regional variation across an object; Phase Congruency (PC) (Kovesi, 1999), which is more akin to describing the edge information present in an image, however, unlike traditional edge detectors, PC has the added advantage of incorporating scale information into its calculation, thus tends to highlight significant features that the HVS is thought to respond to; Position (XY), which simply relate to the positional co-ordinates (x,y) of pixels in the image plane, considered here such that partitioning across position will give formed clusters a sense of proximity within the image.

Of course, in reality it is highly likely that there are many additional features and factors that contribute to object discrimination as performed by the HVS.

We also note that, in this application, although we primarily need to segment 3D data obtained from the confocal microscope, we restrict our experiments for the current work to the 2D segmentation of an individual slice, so that the significance of these features with respect to one another (as fused through the SOTM), can be assessed, along with the impact of the proposed refinements to the SOTM model.

6. Results and discussion

Experiments were conducted on individual 2D confocal image slices from a sample of FITC-stained biofilm. Of the slices, one was chosen for this presentation (the right-hand image from Fig. 1), such that adequate comparison could be drawn as to the impact of feature adjustment on the resulting SOTM segmentation. A number of experiments were conducted, and in each, constraints were imposed both in the sets of features to be fused by the SOTM, and by imposing a bias or weighting to the competitive search phase of SOTM learning (step (c) in the SOTM algorithm). Three main dynamics were explored.

6.1. Effects of feature space definition

The first experiments explored the dynamic of feature space definition. In this test, our goal was to consider how the SOTM responded when features were added to the feature space description, one by one. With each new definition, we evaluate the impact on segmentation. A 64 neuron SOTM was utilised in this process. The new stop criterion was not enforced so that the network could grow to the same number of neurons for each experiment.

Feature spaces (*FSi*), were defined as follows, with reference to Section 5. The segmentation results are presented in Fig. 4 (a) $FS_1 = (GL)$; (b) $FS_2 = (GL, GL_{dsk})$; (c) $FS_3 = (GL, GL_{dsk}, PC)$; (d) $FS_4 = (GL, GL_{dsk}, PC, XY)$. We consider both X and Y to be of equal significance, and to be used in unison to represent position of a pixel. All features remained unbiased throughout learning, and were normalised to eliminate any possibility of natural ranges in the data adversely affecting the outcome of the segmentation. Two images were obtained for each case, however we omit the segmented dataset labelled with the GL component of the classifying neuron. The reader is directed to Fig. 5(a) for reference. Since the analysis is subjective, we present an image of the segmentation with a randomly colour-labelled version, such that any assessment of region/object localisation is not biased by our own tendency to segment the resulting GL images naturally using our own built in HVS.[1]

As can be seen from the labelled images on the right, both (a) and (b) demonstrate little in the way of a contiguous region (connected set of labels), with the exception of regions of extreme uniformity in GL existing in the background. With PC, in (c), we start to observe some mild regional associations. The overall segmentation however is still rather contaminated by the signal based pixel statistics, each competing (in an equal fashion) for the right to represent the input space. If one considers the PC feature, in being more representative of significant features and edge information, one might assume that a large dynamic range of PC values might exist, but over a more limited expanse of the input data set. As such the network would tend to encounter large differences in PC more often, thus partition in this dimension more readily. Even with this though, the majority of the labelled segmentation appears quite random and with little grouping across regions.

When adding position to the mix, the SOTM appears to immediately discover more regional based associations.

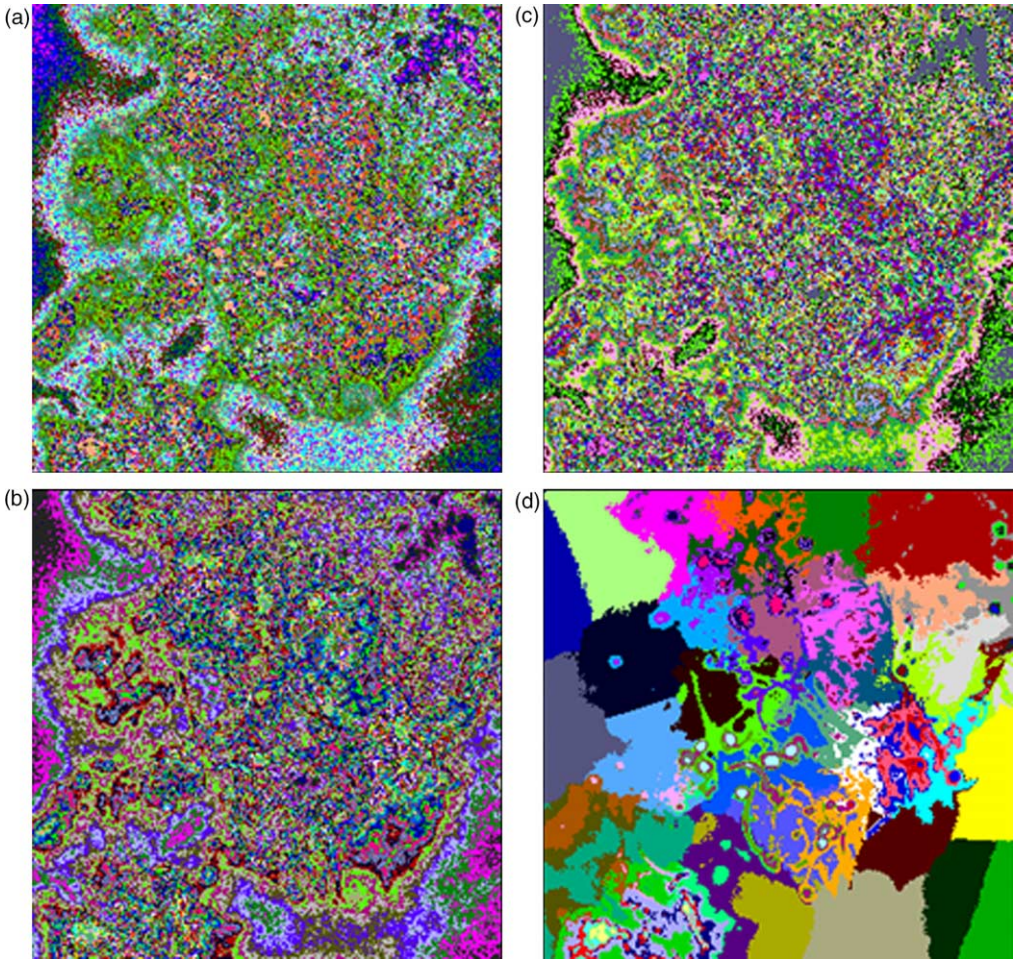

Fig. 4. (a) top left: GL only, (b) bottom left: GL,GL_{dsk} (c) top right: GL,GL_{dsk},PC (d) bottom right: GL,GL_{dsk},PC,XY.

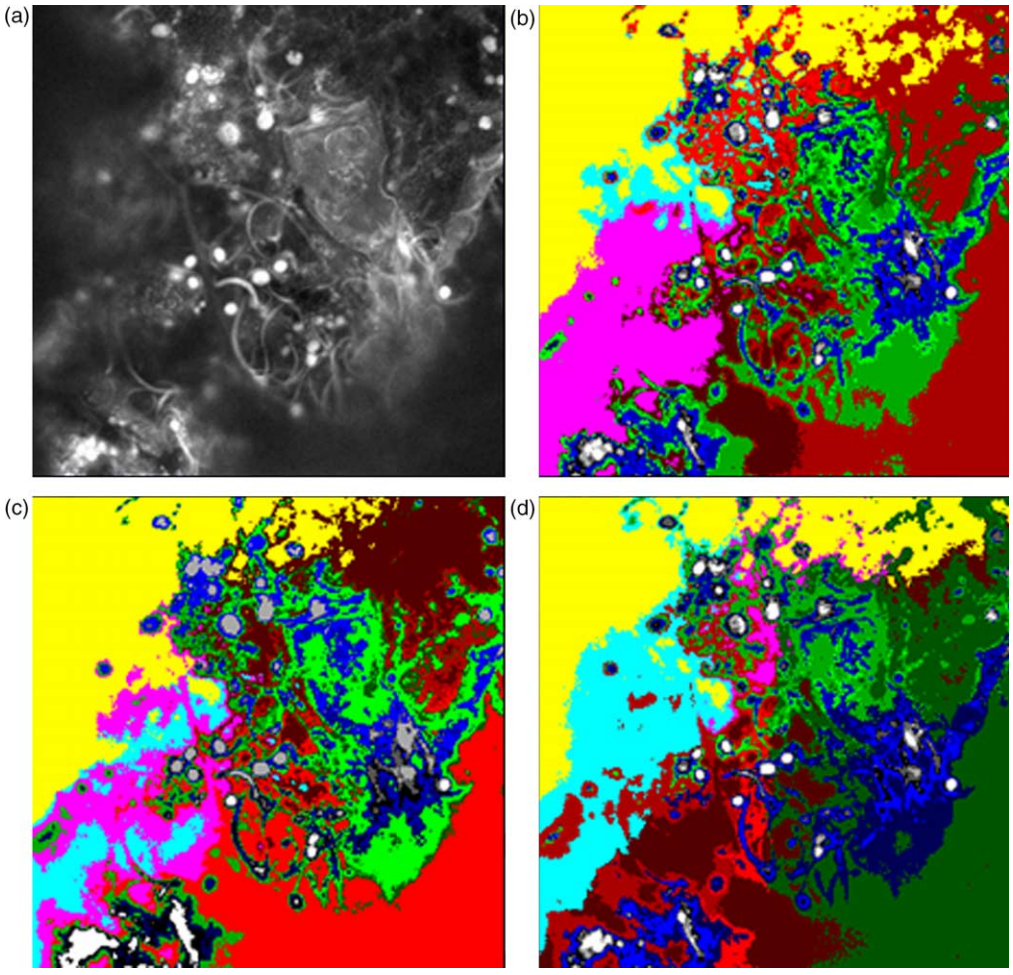

Fig. 5. 16 class SOTM segmentation on all features (a) top left (original slice-for reference), (b) top right: $\beta xy = 0.01$, (c) bottom left: $\beta xy = 0.02$, (d) bottom right: $\beta xy = 0.1$.

Whilst not delineating all structures per se (limited by its restriction to 64 neurons maximum), a good number of neurons appear to favour a relatively uniform partitioning of *XY* space. One would assume that the same occurs in the other dimensions (based on the relative density of information in those dimensions). In the *XY* plane, the density is always uniform, as the *X* and *Y* vectors carry no information on their own. The partitioning in this plane however, allows for neurons to separate and yet still learn from GL regional information. This could offer an insight into how separate notions of objects may be formed within the memory of the SOTM, due to their separation in proximity. At any rate, it seems clear that there is strong interplay between the knowledge of signal characteristics and proximity, and the SOTM demonstrates potential in differentiating this.

6.2. Fixed weighting of feature space

In the second experiment, we consider applying a bias to the signal and spatial components of feature space. In separate runs of a 16 node SOTM, we adjust the significance weighting of *XY* components, to investigate how the segmentation is impacted when the competitive learning process is biased toward different features. At this stage we also draw a comparison between the popular k-means clustering method (Therrien, 1989) and the SOTM. The results are presented in Fig. 5(b)–(d). In all cases: GL, GL_{dsk} & PC remain unbiased (weighting of 1), whilst *XY* is changed. In (b), the weighting for *XY* is 0.01, in (c) 0.02, and (d) 0.1. A high ratio of Signal vs. *XY* weightings favour a relatively weaker partitioning in the *XY* plane. Thus the number of neurons split across (GL,GL_{dsk},PC) will tend to outnumber those split across *XY*. Sensitivity of the SOTM to proximity is then considered via a range of fixed *XY* weightings.

From the segmentation results in Fig. 5, we notice that, as expected, significantly fewer neurons are dedicated to the representation of background as compared with Fig. 4(d), thereby freeing up neurons to capture foreground regions, where microbial material exists (i.e. regions of high signal activity). The extent of these regions is reduced or 'tightened' as we increase the significance of *XY* features (b)–(d), whilst the range of intensities captured by each

neuron (region) is extended or 'relaxed'. From this it becomes evident that there is an inherent trade-off between signal and proximal resolution as we bias the significance ratio between signal and position.

It is at this point that we compare the SOTM's performance with a similar clustering technique: *k*-means. The features were weighted as for Fig. 5(d), however the SOTM was allowed to decide on the appropriate number of clusters by itself. A total of 31 clusters (or class centres was discovered). We then ran the k-means algorithm imposing this limit of 31 clusters, such that we might get a direct indication of performance. We also used a randomly generated, but ordered colour-map, so that the labels thus discovered by each algorithm (also ordered in terms of their greyscale property), might be compared. What we find from Fig. 6 is that there is quite a marked clarity exhibited through the SOTM result. Clusters discovered tend to form well packed groupings of similar signal characteristics. The red and purple sets of nodules (large bacteria clumps), for instance are quite pronounced and grouped into two localised regions, as opposed to k-means case. This is likely due to the systematic way in which the SOTM parses the feature space. Partitioning is biased toward the selective allocation of new classes (neurons) with the occurrence of either slight variations in signal or large variations in the proximity of similar signals.

The k-means approach does not share this flexibility, as it essentially begins with a choice of random samples from across the entire feature space, attempting to then adjust the centres from what may in fact be ill-fated initial positions.

Perceptually, the SOTM gives a much clearer result, and appears to offer a more efficient representation across feature space. The higher the feature space dimension grows in any one application, the more important this property will become. Automatic selection of class numbers is also a major advantage of the SOTM here.

6.3. Dynamic weighting of feature space

In the final experiment, we investigate what would happen if there is compromise between features over the course of learning. Using the modifications discussed in Section 4.2, we show results for the dynamic weight adjustment of *XY* features, decreasing or increasing their role in the search phase of competitive learning, over time. Fig. 7 shows the results and experimental conditions.

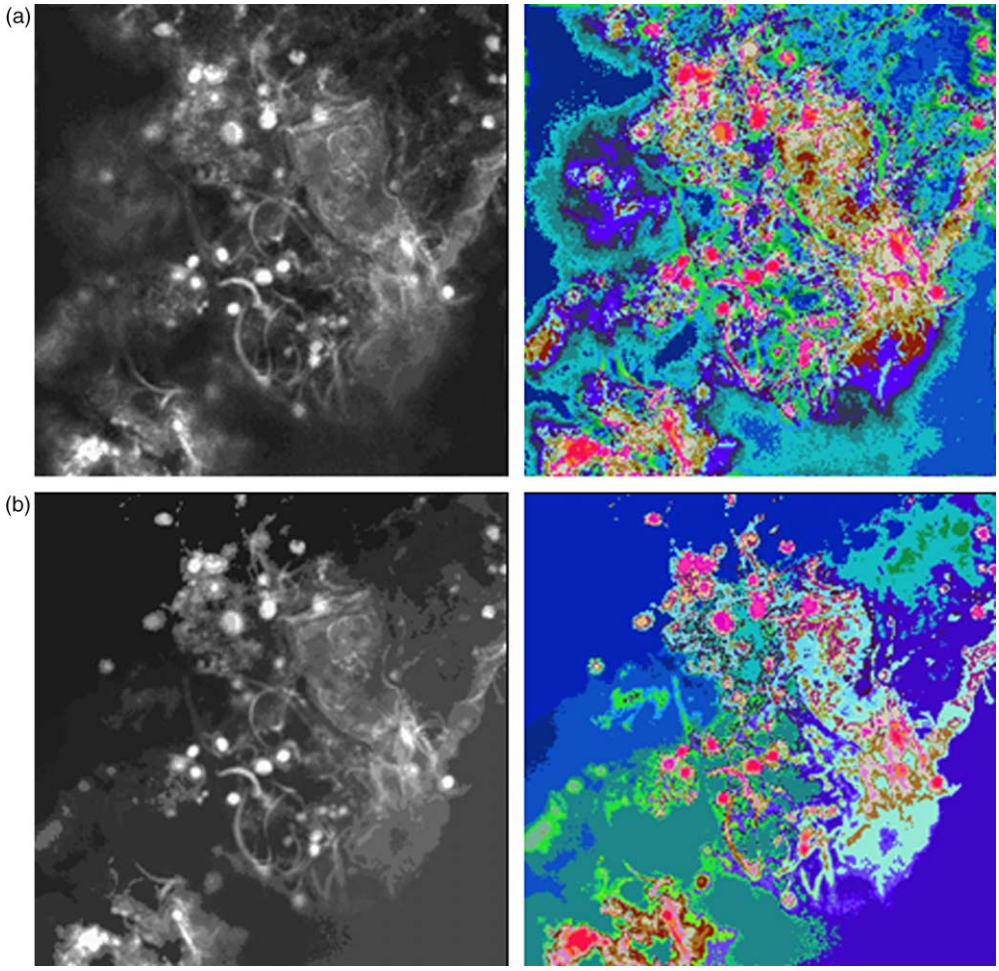

Fig. 6. (a) top: *k*-means, 31 centres, (b) bottom: SOTM, 31 centres (self discovered).

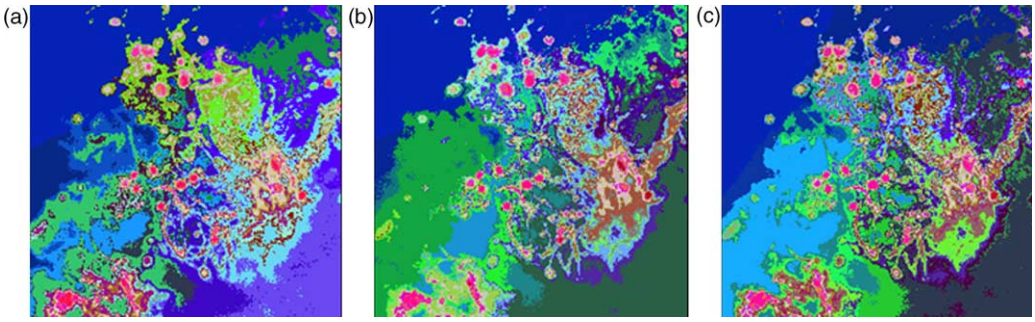

Fig. 7. (a) left: 28 centres, βxy decayed from 1 over duration 1*Ns, H(∞)=0.3 (b) middle: 47 centres, βxy decayed from 1 over 2*Ns, H(∞)=0.2 (c) right: 42 centres, βxy grown from 0, H(∞)=0.2.

We also reduce the minimum limit on $H(t)$ for cases (b) and (c) to see what happens if more neurons are allowed to generate.

In each case (a) and (b), we begin with an initial XY weight of 1, which then decays to zero over two different periods (denoted by multiples of Ns: the number of samples in the image data). In (a), more vigorous partitioning occurs in XY whilst it is as significant as the GL features, thus tight regions are captured as in Fig. 4(d). This is somewhat offset though, by the large initial values of $H(t)$, which force partitioning in XY to only occur across well separated data. In essence, if two data points are encountered that have similar signal properties, then their separation must be large before if they are to be considered as different. As the XY significance switches to zero, signal features now dominate and neurons generated tend to be pushed toward capturing signal variations. Unlike Fig. 4(d), background regions not having much variation in signal tend not to be further partitioned as the constraint on XY is relaxed, whilst biomass regions (foreground) are. The result is more detail in the foreground (which is what we prefer to delineate). In (b), we relax the XY weighting more slowly. The danger here is that as $H(t)$ decreases, more of the background partitions may in fact be split, whilst the same is true for foreground. Such is a trade-off that may need to be addressed through some mechanism to reduce the impact of background partitions. Some domain knowledge can be used here (background should have the lowest intensities). Fig. 6(c) shows the opposite case: growth of the XY weighting. This pushes neurons to represent signal early, then split across XY later, the same danger thus exists if the weight grows too fast, before $H(t)$ is limited.

7. Conclusions

The flexibility and efficiency of the SOTM to adaptively build a representation of the input feature space has revealed interesting properties via dynamic adjustment of feature significance in the competitive search phase of learning. In this early stage of testing, it is clear that there is an important interplay between positional and graylevel characteristics when building a representation of an ‘object’ or ‘region’ in an image with the SOTM. Evidence suggests that the SOTM can adaptively warp its topology from the traditional means squared sense, into forms that better reflect the types of segmentations typical of the human visual system. Continued investigation and understanding of such properties will be crucial in such applications where segmentation is a difficult task-such as in the extraction of relevant information from micro-biological confocal images.

Acknowledgements

The authors (Steven Liss and Ling Guan) wish to acknowledge the financial support from the Natural Sciences and Engineering Council of Canada. Matthew Kyan gratefully acknowledges the APA scholarship from the Australian Research Council. Images of biofilms were provided by Xuetang Zhu, PhD Candidate, Department of Chemical Engineering and Applied Chemistry, University of Toronto, supervised by Steven Liss.

References

Baveye, P. (2002). Comment on ‘Evaluation of biofilm image thresholding methods’. *Water Research*, *36*, 805–806.

Besag, J. (1986). On the statistical analysis of dirty pictures (with discussion). *Journal of Royal Statistical Society. Series B*, *48*, 259–302.

Beyenal, H., Lewandowksi, Z., & Harkin, G. (2004). Quantifying biofilm structure: Facts and fiction. *Biofouling*, *20*(1), 1–23.

Kittler, J., & Illingworth, J. (1986). Minimum error thresholding. *Pattern Recognition*, *19*, 41–47.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*, 59–69.

Kong H. S. (1998). The self-organising Tree map, and its applications in digital image processing. PhD Thesis. University of Sydney, Australia.

Kovesi, P. (1999). Image features from phase congruency. Viere. *Journal of Computer Vision Research*, *1*(3), 1–26.

Kyan, M., Guan, L., & Liss, S. (2005). *Dynamic feature fusion in the self organising Tree map—applied to the segmentation of biofilm images Proceedings of the International Joint Conference on Neural Networks. July 31–August 4, 2005, Montreal, Canada.*

Liss, S. N. (2002). Microbial flocs suspended biofilms. In G. Sutton (Ed.), *Encyclopedia of environ. microbiology* (pp. 2000–2012). New York: Wiley.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions On Systems Man and Cybernetics*, *9*(1), 62–66.

Palmer, R. J., & Sternberg, C. (1999). Modern microscopy in biofilm research: confocal microscopy and other approaches. *Current Opinion in Biotechnology*, *10*(3), 263–268.

Randall, J., Guan, L., Zhang, X., & Li, W. (1999). *Investigations of the self-organising tree map Proceedings of ICONIP' 99* pp. 724–728.

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

Therrien, C. W. (1989). *Decision estimation and classification—An introduction to pattern recognition and related topics*. New York: Wiley pp. 215–227.

Tolle, .C. (2003). Suboptimal minimum cluster volume cover-based method for measuring fractal dimension. *IEEE Transactions On Pattern Analysis and Machine Intelligence*, *25*(1), 32–41.

Wilkinson, M. H. F. (1998). Automated and manual segmentation techniques in image analysis of microbes. In M. H. F. Wilkinson, & F. Schut (Eds.), *Imaging, morphometry, fluorometry and motility techniques and applications*. New York: Wiley.

Yang, X., Beyenal, H., Harkin, G., & Lewandowski, Z. (2001). Evaluation of biofilm image thresholding methods. *Water Research*, *35*(5), 1149–1158.