Segmentation of Dense 2D Bacilli Populations

Pascal Vallotton
CSIRO, Division of Mathematics, Informatics, and Statistics
Sydney, Australia
pascal.vallotton@csiro.au

Lisa Mililli
Department of Mechanical Engineering
Monash University
Melbourne, Australia

Lynne Turnbull, Cynthia Whitchurch
Institute for the Biotechnology of Infectious Diseases
UTS
Sydney, Australia

Abstract— **Bacteria outnumber all other known organisms by far so there is considerable interest in characterizing them in detail and in measuring their diversity, evolution, and dynamics. Here, we present a system capable of identifying rod-like bacteria (bacilli) correctly in high resolution phase contrast images. We use a probabilistic model together with several purpose-designed image features in order to split bacteria at the septum consistently. Our method commits less than 1% error on test images. Our method should also be applicable to study dense 2D systems composed of elongated elements, such as some viruses, molecules, parasites (plasmodium, euglena), diatoms, and crystals.**

Keywords- Bacteria; motility; segmentation; image analysis; Bayesian; probability; phenotype

I. INTRODUCTION

The ability to image and characterize single cells is becoming increasingly important in biology as it allows addressing questions about population genetics, biodiversity, aging, and noise in biological systems at unprecedented detail [1]. Identifying and tracking single cells is also fundamental to understanding how motility is involved in diseases such as cancer and bacterial infections.

When observed under natural conditions, bacterial cells in a dense population are very difficult to distinguish from each other as they tend to aggregate and are multi-layered. Thus, experimental systems for observing bacterial dynamics in controlled conditions have been developed, with particular attention to the ability of observing surface colonization; an important mechanism underlying the pathogenicity of opportunistic bacteria [2]. Using this technology, a trained operator can attempt to manually count and characterize bacteria. However, this represents a daunting task as there may be thousands of cells in a single frame (Fig. 1). Thus, many researchers have attempted to automate the process. For example, Wang describes a fairly generic system capable of segmenting and tracking cells of various origins [3]. As it is designed to be so generally applicable, it is also relatively slow. Other researchers focus on characterizing bacteria in more dilute environments [4, 5]. Other researchers do not identify individual bacteria but work directly with the image velocity field [6]. Vallotton *et al.* describe a system capable of segmenting and tracking thousands of bacteria in dense populations but they do not address adequately the problem of cell division – thus frequently treating two or more cells as a single cell [7].

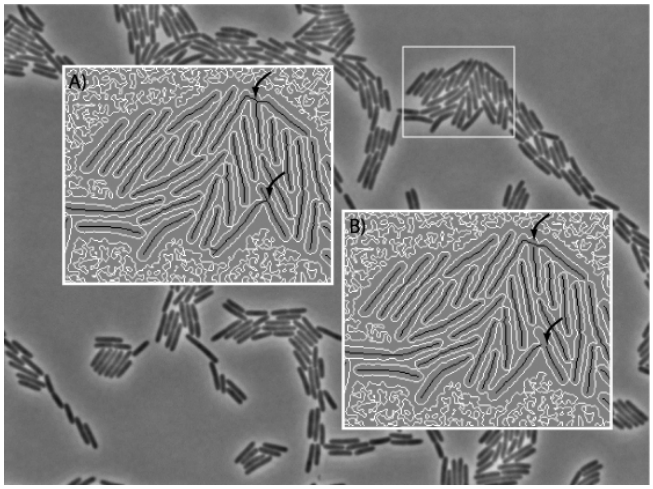

Figure 1. Phase contrast microscopy image of Pseudonoma aeruginosa confined in two dimensions. Inset A) shows the regions as defined by the edge detection step together with the morphological skeletons. Inset B shows the skeleton after crossing points have been eliminated. Skeletons spanning several bacteria necessitate further subdivisions as described in the text.

Bacteria possess characteristic shapes, predominantly rod-like (bacilli), spherical (cocci), or helical (spirillum). Such shape constraints can be exploited for designing optimized image analysis solutions. In this manuscript, we deal exclusively with bacilli, for which most shape information is contained in the probability distribution characterizing their length. We describe how this can be exploited to segment bacteria arranged in tandem. A full scheme that exploits the segmentation method outlined in

978-0-7695-4271-3/10 $26.00 © 2010 IEEE
DOI 10.1109/DICTA.2010.23

this manuscript to analyze bacterial dynamics during surface colonization is under development.

II. MATERIALS AND METHODS

A. Sample preparation and Imaging

Microscopy of sub-surface colony expansion was performed using a modified dipped-slide assay [2]. Briefly, glass slides were flame-sterilized and three were placed side by side in a 90 mm petri dish. 5 mL of nutrient medium (8 g/L tryptone, 4 g/L yeast extract, 4 g/L NaCl) solidified with GelGro (ICN) gellan gum (8 g/L) instead of agar for greater optical clarity) was poured onto the 3 slides and allowed to set. The slides were then separated by gently cutting the solidified media. Each slide was inoculated with bacteria onto the solidified nutrient media. A sterilized glass coverslip was placed onto the bacteria and the slide incubated for 4 hr at 37^{o}C in humidified conditions. Colony expansion at the interstitial surface was then observed through the coverslip using phase contrast microscopy with a Nikon Ti inverted microscope fitted with a 100x oil objective (1.4 NA) and an incubation chamber to maintain the temperature at 37^{o}C during image acquisition. The large NA and high magnification objective enabled high resolution imaging of individual bacterial cells in the expanding colony. Under phase contrast, bacteria appeared dark on a light background, such that they can be distinguished from each other by eye (Fig.1). Pixel size in the object space was 60 nm.

B. Image Analysis

1) Boundary detection

The outer limit of bacilli defines elongated and closed regions. Hence, it is appropriate to select an edge detector that guarantees closed contours. This is the case of the Marr Hildreth edge detector defined by the zero level set of the image Laplacian [8]:

$$\Delta I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} = 0$$

Each closed contour defines a region containing one or more bacteria, most often arranged in tandem (we use the words “region”, “segment”, “fragment”, and “closed contour” interchangeably). Exceptionally, the region is not linear because some bacteria abut onto each other, forming a crossing point (Fig. 1). We replicated the method of Vallotton et al. to deal with these crossing points as they must be eliminated for a physically acceptable segmentation. First we produced the skeleton by morphological thinning of our regions [9]. Then, we cut the skeleton at crossing points, gluing it back together such that only the two segments that possess the most similar orientations are joined [7].

This still leaves large assemblies of bacteria as fragments comprising from 2 to 4 bacteria in tandem. It is particularly important to split such assemblies if one aims to track bacteria over time, because treating bacteria alternately as a singlet or as a doublet creates rugged trajectories, or can even break them. This is also important for avoiding false positives when screening for mutant bacterial phenotypes.

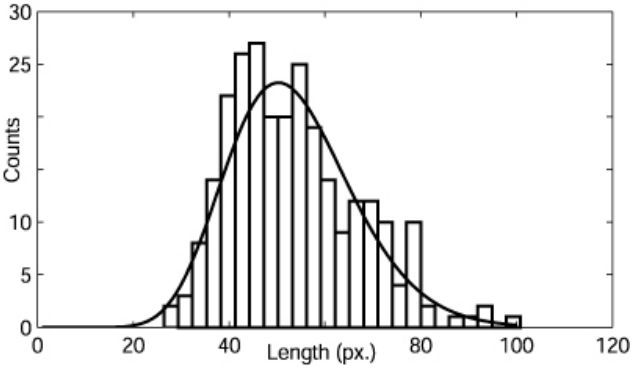

Figure 2. The bacterial length l is approximately modeled by a gamma distribution $\Gamma_{16.41,3.26}(l)$. The histogram was obtained by manual image measurements.

2) Splitting adjoining bacteria

a) General remarks

Bacteria tend to adhere to each other via their extremities - an association presumably mediated by pili - long fibers implicated in motility and projecting from the bacterial poles [10]. A tandem configuration also results from bacterial division, which proceeds by invagination at the septum. Following division, bacteria tend to remain connected to their sister cell for extended periods.

b) Selective features for detecting splitting

The image intensity increases at the septum during cell division as the daughter bacteria move away from each other, opening a gap. Thus, candidate splitting points are also local intensity maxima along the bacterial skeleton (Fig. 1).

The image contrast at a particular candidate point is the first feature that we exploit to classify candidate cuts (Fig. 3B). In order to measure the contrast, we perform a morphological top-hat with a disk structuring element of diameter 13 pixels [9].

Another landmark of bacterial division is provided by a constriction at the septum. We measure this narrowing via the Euclidian distance transform performed on the binary regions defined by the edge detection step (Fig. 3A).

As additional cue for bacterial division, we also estimated the angle formed by the bacterial segments on each side of a candidate cutting point. Each candidate point is considered in turn. The two neighbor candidates on both sides define the angle in question. We found this more reliable than using the candidate splitting point as the angle apex because abutting bacteria are not always perfectly aligned.

Our third, and probably more interesting feature, is provided by a probabilistic model P(x|L) which estimates the likelihood of a cut at particular position x along a bacterial skeleton of length L (Fig. 3C). This is described in the next section.

All four image features were normalized between 0 and 1 and multiplied to quantify the total evidence for a cut. A threshold is finally selected depending on the application. We used 0.1 as a threshold in this manuscript.

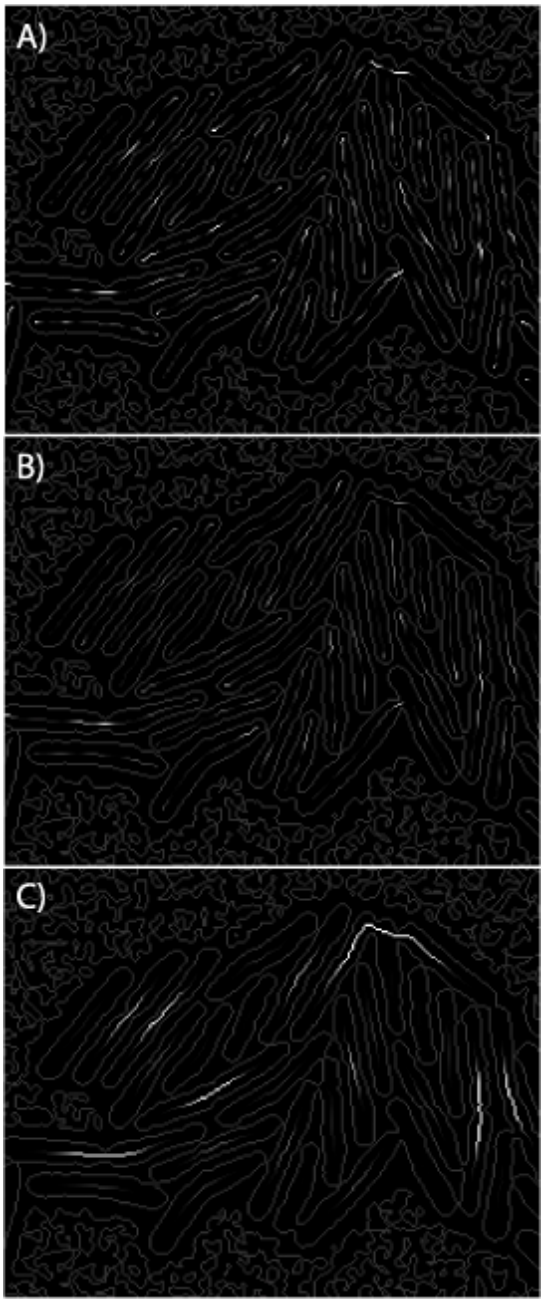

Figure 3. Image features used for increasing the reliability of bacterial region splitting. Panel A) maps the local narrowing measured via the distance transform. Panel B) maps the local intensity increase along the skeleton. Panel C) maps the splitting probability. The map of the angular feature is not shown. All four features, normalized between 0 and 1, are multiplied together to estimate the evidence for a cut.

c) Modeling the probability of true cuts in bacterial skeletons

We found that the bacterial length l is well approximated by a gamma distribution $\Gamma_{\alpha,\beta}(l)$ (Fig. 2). A region of length well below the average length should probably not be divided. Likewise, a region twice the average length of a bacterium should be examined closely for evidence of a potential cut.

Let $P(x_1)$ stand for the probability of a cut at a distance x_1 from an extremity. Confronted with a region of length L, we compute

$$P(x_1 \mid L) = \frac{P(x_1, L)}{P(L)} = \frac{P(x_1)P(L-x_1)}{\int_{l=0}^{L} P(x)P(L-x)dx} \quad (1)$$

The expression in the denominator is the distribution of the sum of two random variables [11]. The expression for the numerator assumes that the lengths of two bacteria within the same segment are independent of each other. This is only an approximation as bacterial division tends to produce daughter cells of approximately equal size. However, as bacteria without filial relation often come in tandem, we decided not to take this effect into account.

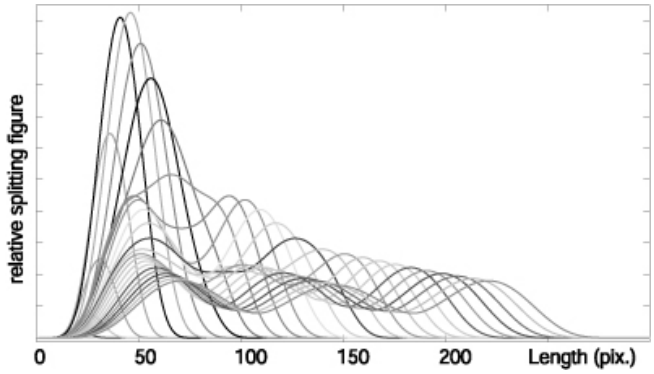

Figure 4. Relative splitting probability factor as a function of the length of fragments. These curves are obtained from the cut probability P(x|L) as described in the text. These curves are mapped onto the skeleton in Figure 3C.

If x is distributed according to $\Gamma_{\alpha_1,\beta}$ and y is distributed according to $\Gamma_{\alpha_2,\beta}$, then x+y is distributed according to $\lambda_{\alpha_1+\alpha_2,\beta}$ [11].

Thus:

$$P(x_1 \mid L) = \frac{\Gamma_{\alpha,\beta}(x_1)\Gamma_{\alpha,\beta}(L-x_1)}{\Gamma_{2\alpha,\beta}(L)} \quad (2)$$

It is not exceptional to find three bacteria in a single region. One would favor cutting such a region twice, at positions x_1, and x_2, such that x_1, x_2-x_1, and L-x_2 are approximately equal. Let $P(x_1, x_2 \mid L)$ denote the probability of the corresponding configuration. By marginalizing on x_2,

$$P(x_1 \mid L) = \int_{x_1}^{L} P(x_1, x_2 \mid L)dx_2 \quad (3)$$

Also, by definition of a conditional probability:

$$P(x_1, x_2 \mid L) = \frac{P(x_1, x_2, L)}{P(L)} \tag{4}$$

Assuming as above that the length of bacteria are independent of each other in a given region:

$$P(x_1, x_2, L) = P(x_1)P(x_2 - x_1)P(L - x_2) \tag{5}$$

As before, we can obtain the distribution for the sum of three random variables by convolution (x_1 and x_2 always come in the order stated):

$$P(L) = \int_0^L P(l_1)dl_1 \int_0^{L-l_1} P(l_2)P(L - l_1 - l_2)dl_2 \tag{6}$$

Exploiting again the additive property of gamma distributions, we obtain:

$$P(x_1 \mid L) = \frac{P(x_1)\int_{x_1}^{L} P(x_2 - x_1)P(L - x_2)dx_2}{\Gamma_{3\alpha,\beta}(L)} \tag{7}$$

By substitution,

$$P(x_1 \mid L) = \frac{\Gamma_{\alpha,\beta}(x_1)\Gamma_{2\alpha,\beta}(L - x_1)}{\Gamma_{3\alpha,\beta}(L)} \tag{8}$$

$P(x_2|L)$ is obtained by a symmetry argument or by marginalizing on x_2.

The case of four bacteria making up a single region is very similar:

$$P(x_1 \mid L) = \frac{\Gamma_{\alpha,\beta}(x_1)\Gamma_{3\alpha,\beta}(L - x_1)}{\Gamma_{4\alpha,\beta}(L)} \tag{9}$$

$P(x_2|L)$ is obtained by marginalizing on x_1 and x_3:

$$P(x_2 \mid L) = \frac{\Gamma_{2\alpha,\beta}(L - x_2)\Gamma_{2\alpha,\beta}(x_2)}{\Gamma_{4\alpha,\beta}(L)} \tag{10}$$

Although we never found in excess of 4 bacteria per region in our application, some species tend to form very long strings of bacteria, so it is worth stating the generalization to N cuts:

$$P(x_k \mid L, N) = \frac{\Gamma_{k\alpha,\beta}(x)\Gamma_{(N-k)\alpha,\beta}(L - x)}{\Gamma_{N\alpha,\beta}(L)} \tag{11}$$

Up to now, the number of cuts was always implicit (we wrote $P(x_1,x_2|L)$ rather than $P(x_1,x_2|L,n=3)$. The probability of a cut is given by marginalizing over the number of bacteria in a region (1, 2, 3 or 4):

$$P(x \mid L) = \frac{P(x, L)}{P(L)} = \frac{\sum_{i=1}^{4} P(x, L \mid n = i)P(n = i)}{P(L)} \tag{12}$$

Likewise, for the denominator:

$$P(L) = \sum_{k=1}^{4} P(L \mid k)P(k) = \sum_{k=1}^{4} \Gamma_{k\alpha,\beta}(L)P(k) \tag{13}$$

Finally, as all cut indices do contribute to the probability of a cut at a particular position:

$$P(x, L \mid n = i) = \sum_{k=1}^{i} P(x_k, L \mid n = i) \tag{14}$$

We measured on a sample of approximately 1000 cells that $P(n=4) \simeq 0.01$, $P(n=3) \simeq 0.01$, and $P(n=2) \simeq 0.03$.

All calculations were performed under Matlab™. In order to obtain an image feature from P(x|L), the latter was scaled by L because as a normalized probability distribution, P has a decreasing amplitude when the segment becomes longer. Finally, the square root of P was used, rather than P, which we found useful when dealing with some exceptional configurations of bacteria. The final response functions are plotted in Figure 4 as a function of the segment length. To use these response functions as image features, we map them onto the skeletons, as shown in Figure 3C.

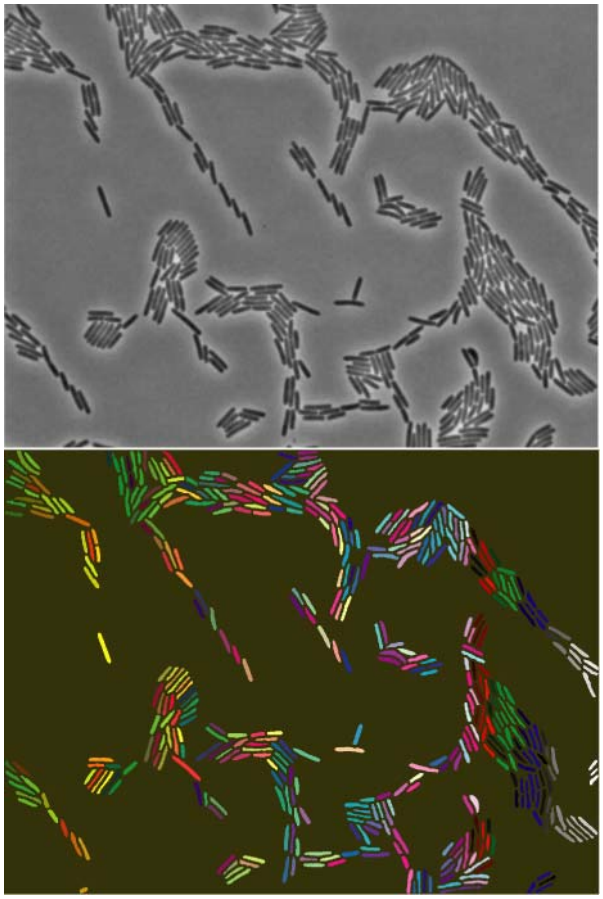

Figure 5. Final segmentation results. In excess of 500 cells can be counted in this frame, all of them correctly segmented. Note that segments touching the image border have been removed.

III. RESULTS

The candidate cutting points above the selected threshold are removed from the skeleton, prior to running the watershed transform, seeded by the broken skeletons [9]. Also, spurious objects in the background are removed based on contrast and shape criteria. The final results for a particular frame are shown in Figure 5. These results appear faultless based on the observation of a single frame. A more difficult problem, which we do not address here, is to guarantee the consistency of the splitting decision across multiple frames.

The images were of size 1040*1392 pixels and were converted to type double for convenience. Processing speed was less than 10 seconds per frame on a contemporary laptop

(Dell latitude™). In terms of bacterial counting, this represents only a modest improvement over manual counting (it is possible to count all bacteria within a few minutes). However, the main advantage of the system is that it provides reliable information on the shape of bacteria as well. For this application, we estimate that the improvement could be over a hundred fold as it takes a few seconds to manually trace the outline of a single bacterium with decent precision.

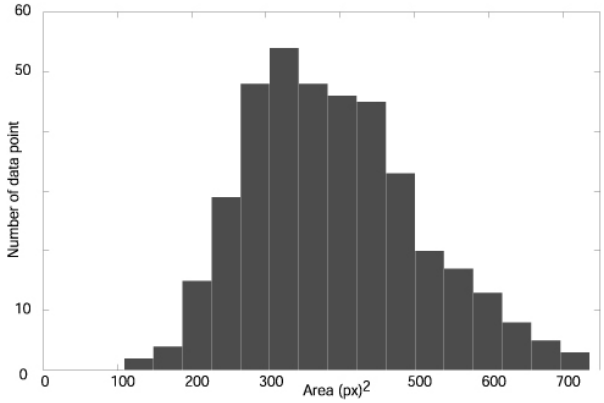

Figure 6. Bacterial segmentation allows accessing the distribution of phenotypic traits. Here, the size histogram is shown.

IV. CONCLUSION

An effective method for segmenting bacteria was presented. The combination of multiple image features considered simultaneously ensures that cutting decisions are made reliably. The use of the Marr Hildreth edge detector proved superior to the Canny detector used by Vallotton *et al.*, such that we did not find that a linear feature detector had to be added to the image analysis workflow to obtain good results. Besides guaranteeing closed contours, Marr's detector possesses optimal edge integration properties, which may contribute to it's suitability for our application [8]. It may be helpful to customize the edge detection step itself to include a priori knowledge of the bacterial shape.

One immediate application of our research is the characterization of bacterial cells on very large scale for studying phenotypic variability and forward genetics. For example, the distribution of bacterial size in Figure 5 could serve as the starting point for selecting large bacteria or bacteria presenting delayed division.

A more challenging application of our work is to understand the dynamics of bacterial populations spreading over surfaces, such as biological tissues. We now can hope to address quantitatively questions relevant to bacterial memory, lattice formation, cell-cell interactions and signaling [12-16]. This will require segmenting bacteria consistently over thousands of image frames.

ACKNOWLEDGMENT

We thank Mitchell Buckley, David Lovell, and two anonymous reviewers for providing useful input.

REFERENCES

[1] C. L. Woldringh, Is Escherichia coli getting old? Bioessays, vol. 27, pp. 770-774, Aug 2005.

[2] A. B. Semmler, C. B. Whitchurch, and J. S. Mattick, A re-examination of twitching motility in Pseudomonas aeruginosa Microbiology, vol. 145, pp. 2863-73, 1999.

[3] Q. L. Wang, J. Niemi, C. M. Tan, L. C. You, and M. West, Image Segmentation and Dynamic Lineage Analysis in Single-Cell Fluorescence Microscopy, Cytometry Part A, vol. 77A, pp. 101-110, Jan.

[4] J. Xie, S. Khan, and M. Shan, Automatic Tracking of Escherichia Coli in Phase-Contrast Microscopy Video, IEEE Transactions on Biomedical Engineering, vol. 56, pp. 390-399, Feb 2009.

[5] K. Hoshi and R. Shingai, Computer-driven automatic identification of locomotion states in Caenorhabditis elegans, Journal of Neuroscience Methods, vol. 157, pp. 355-363, Oct 2006.

[6] E. B. Steager, C. B. Kim, and M. J. Kim, Dynamics of pattern formation in bacterial swarms, Physics of Fluids, vol. 20, p. 5, Jul 2008.

[7] P. Vallotton, C. Sun, D. Wang, P. Ranganathan, L. Turnbull, and C. Whitchurch, Segmentation and tracking of individual Pseudomonas aeruginosa bacteria in dense populations of motile cells. In Image and Vision Computing New Zealand (IVCNZ) Wellington, New Zealand, 2009.

[8] R. Kimmel and A. M. Bruckstein, Regularized Laplacian zero crossings as optimal edge integrators, International Journal of Computer Vision, vol. 53, pp. 225-243, Jul-Aug 2003.

[9] P. Soille, Morphological Image Analysis, 2 ed, Springer, 2004.

[10] L. L. Burrows, Weapons of mass retraction, Mol Microbiol. vol. 57, pp. 878-88. 2005.

[11] S. Ross, Introduction to Probability Models: Academic Press, 2006.

[12] E. Ben-Jacob, Learning from bacteria about natural information processing, Ann N Y Acad Sci., vol. 1178, pp. 78-90. 2009.

[13] N. Chia, C. R. Woese, and N. Goldenfeld, A collective mechanism for phase variation in biofilms, Proc. Natl. Acad. Sci. U S A. vol. 105, pp. 14597-602. Epub 2008 Sep 17. 2008.

[14] G. Ebersbach and C. Jacobs-Wagner, Exploration into the spatial and temporal mechanisms of bacterial polarity, Trends Microbiol. vol. 15, pp. 101-8. Epub 2007 Feb 1, 2007.

[15] O. A. Igoshin, A. Mogilner, R. D. Welch, D. Kaiser, and G. Oster, Pattern formation and traveling waves in myxobacteria: theory and modeling, Proc. Natl. Acad. Sci. U S A., vol. 98, pp. 14913-8. 2001.

[16] W. Margolin, Sculpting the bacterial cell, Curr Biol., vol. 19, pp. R812-22. 2009.