Automatic Identification and Classification of Bacilli Bacterial Cell Growth Phases

P.S. Hiremath
Department of Computer Science,
Gulbarga University, Gulbarga, Karnataka, India

Parashuram Bannigidad
Department of Computer Science,
Gulbarga University, Gulbarga, Karnataka, India

ABSTRACT
A major challenge in microbial ecology is to develop reliable and facile methods of computer assisted microscopy that can analyze digital images of complex microbial communities at single cell resolution, and compute useful quantitative characteristics of their organization and structure without cultivation. The objective of the present study is to develop an automatic tool to identify and classify the bacterial growth phases of bacilli cells in digital microscopic cell images. Geometric features are used to identify the different growth phases of bacilli bacterial cells, namely, normal, grownup and about-to-divide. The current methods rely on the subjective reading of profiles by a human expert based on the various manual staining methods. In this paper, we propose a method for bacterial cell classification based on their different growth phases by segmenting digital bacilli bacterial cell images and extracting geometric features for cell growth phase identification and classification using 3σ classifier, k-NN classifier, Neural Network classifiers and Fuzzy classifiers. The experimental results are compared with the manual results obtained by the microbiology expert and demonstrate the efficacy of the proposed method.

Keywords
Cell classification; segmentation; bacterial image analysis; bacilli; cell growth phases; k-NN classifier; Neural Network classifier; 3σ classifier, Fuzzy classifier.

1. INTRODUCTION
Microscopy is one of the most important techniques in microbial ecology, since this is the most direct approach to examine the microbe's world from its own perspective. The value of quantitative microscopy in studies of microbial ecology can be increased even further when used in conjunction with computer-assisted image analysis. There are two main advantages of using digital image processing and pattern recognition techniques in conjunction with microscopy for quantitative studies of microbial ecology. First, automatic image analysis reduces the amount of tedious work with microscopes needed to perform a more accurate quantitative analysis of *in situ* microbial abundance and metabolic activity. Secondly, these techniques provide an important quantitative tool to analyze the structures and spatial features of complex microbial communities *in situ* without cultivation. Five major types of information useful in microbial ecology can be extracted from resolved and segmented microscopical images of growing microbial communities *in situ*. These include recognition of cellular morphological diversity, cell abundance, and spatial, metabolic, and phylogenetic relationships of cells to each other and their surrounding environment.

The process of semi-automatic image analysis of cells to evaluate these aspects of microbial communities can be principally divided into four stages: (i) interactive image acquisition, digitization, and segmentation to locate cells; (ii) automatic measurement to extract features of interest; (iii) classification of different cell growth phases; and (iv) statistical analysis, computations, and interpretation of the data. One of the most important and yet most tedious tasks performed during microscopic analysis of microbial communities is the classification of observed cells into known morphological categories and recognition of new categories as well if new distinct characteristics are captured[16]. A major challenge in microbial ecology is to develop reliable and facile methods of computer assisted microscopy that can analyze digital images of complex microbial communities at single cell resolution, and compute useful quantitative characteristics of their organization and structure without cultivation.

Bacteria are unicellular microscopic organisms which can only be seen through microscope. Bacteria exist in different sizes and shapes and they measure in micro-meter (which is a millionth part of a meter). Bacteria are found everywhere and in all types of environments. There are numerous types of bacteria in the world. Bacteria are mainly classified based on their shapes, biochemistry and staining methods [8]. Recently, along with the morphology, other profiles such as their metabolic activities, conditions required for their growth, biochemical reactions, antigenic properties, and other characteristics are also helpful in classifying the bacteria. However, each type of bacteria has its own characteristics. Most of the bacilli bacteria are characterized by three main different growth phases: normal, grownup and about-to- divide.

The statistical imaging method for automatic identification of bacterial types is proposed by Trattner and Greenspan [13], The artificial neural network approach for bacterial classification has been investigated by Nicolas Blackburn, *et al*. [9]. The data mining techniques are employed for the classification of HEp-2 cells by Petra Perner [11], in which a simple set of shape features are used for classification of bacterial cells. Hiremath and Parashuram [4,5,6] have investigated the automatic classification of bacterial cells and its different growth phases using digital microscopic images using geometric shape features. A computer-aided system for the image analysis of bacterial morphotypes in microbial communities using geometric shape features has been investigated by J. Liu *et al*. [7]. Thomas Posch *et al.*[14] have proposed a new image analysis tool to study biomass and morphotypes of three major bacterioplankton groups in an alpine lake using geometric features. Carolina Wählby *et al*. [2], have investigated algorithms for cytoplasm segmentation of fluorescence labeled cells using statistical analysis techniques based on shape descriptive features.

In this paper, the objective is to propose a method for automatic identification and classification of bacterial bacilli cells in digital microscopic images using geometric features that characterize the different growth phases of bacterial cells using 3σ classifier, k-NN classifier, Neural Network classifiers and Fuzzy classifiers. The experimental results are compared with the manual results obtained by microbiology expert and demonstrate the efficacy of the proposed method.

2. MATERIALS AND METHODS

The spread plate technique is used for the separation of a dilute mixed population of micro-organisms, so that individual colonies can be isolated. Aseptically transfer the scoopful of mixed culture on the Nutrient Agar medium. Spread uniformly with the help of L-shaped spreader on the surface of medium plates. After spreading, incubate at 37°C for 24-48 hours. After incubation, single colonies will appear on the Nutrient Agar media plates. Then pickup the single colony, re-inoculate in LB broth and incubate at 37°C at 120 rpm at each intervals of 4,6,8,12,16 hours and spread the culture on the clean glass slide and go further identification by using staining techniques. A smear of mixed culture bacteria is deposited on a glass slide and thoroughly air-dried. It is stained for 1 min in Crystal Violet solution, 1 min in iodine solution, washed for 20s in ethanol and finally, counterstained with safranin for 1 min. The glass slide is examined under oil immersion at greater than 10000x magnification with direct illumination in a Dialux 20 microscope equipped with a 3 CCD Sony color camera and connected to a PC [1,8]. We have considered 100 color images of each phase of bacilli bacterial cell for present study and these are converted into gray scale images [3].

3. PROPOSED METHOD

The purpose of the automated image analysis of digital bacterial cell image is to identify the different growth phase of bacilli bacteria whether it is normal or grownup or about-to- divide based on their geometric features using different classification techniques, namely, 3σ classifier, k-NN classifier, Neural Network classifiers and Fuzzy classifiers.

Out of many geometric features used by various authors in the literature [3,15], it is observed that there are five geometric features, namely, circularity, compactness, eccentricity, tortuosity and length-width ratio, which provide better classification results. Hence, we have used these five features, which are defined as given below:

Circularity (x_1) : $4\pi(\text{Area})/\text{perimeter}^2$
Compactness(x_2) : A measure of compactness ($\text{Perimeter}^2/4\pi*\text{Area}$)
Eccentricity(x_3) : It is the ratio of the length of the highest chord of the shape to the longest chord perpendicular to it. i.e. $\text{Length}_{major\_axis}/\text{Length}_{minor\_axis}$
Tortuosity(x_4) : Major axis/perimeter
Length-width ratio(x_5): Major axis/minor axis.

Classification rules

3σ classifier

The bacterial cell images generally contain noise, small debris and artifacts depending on the different staining methods. To remove this debris, we have preprocessed the image by applying morphological operations, namely, erosion, reconstruction and dilation. This stage is of high importance in achieving good results in segmentation and further process. The gray scale image of cells is segmented using the adaptive global thresholding, which yields binary image. After labeling the segmented image, the geometric features x_i, i=1,2,..,5, are extracted for each labeled segment. These features are used as a basis for the cell identification and classification. Using the training set of images (with known cell classification), for each feature $x_i^k, i = 1,2,..5$, of k[th] cell phase, we compute the mean and standard deviation σ_i^k of the sampling distribution of the feature values and store them as knowledge base.. In the testing phase, for a given test image, feature values $x_i^{(test)}$ of the segmented regions (cells) are computed and then cell classification is done using the 3σ rule, namely: For a segmented region in the test image, if the feature values $x_i^{(test)}$ lie in the interval $\pm 3\sigma_i^k$, i=1,2,.., 5, then the region is a cell phase of type k. The k=1,2,3 correspond to normal, grownup and about-to-divide phase, respectively.

k-NN classifier

The k-nearest neighbor (k-NN) classification is performed by using a reference data set (training set) which contains both the input (feature set) and the target variables (known cells) and then by comparing the unknown (test data), which contains only the input variables (features) to that reference set. The distance of the unknown to the k nearest neighbors determines its class assignment by either averaging the class numbers of the k nearest reference points or by obtaining a majority vote from them.

Neural Network Classifier

The input layer has 5 neurons and 5 shape features as inputs, and output layer has three output (growth phases, namely, normal, grownup and about-to-divide). The transfer function used is 'tan sigmoidal', training function used is Levenberg-Marquardt back propagation, the weight/bias learning function is 'gradiant descent' function and the performance function is 'mean square error (mse)' which is set to 0.01. In the case of radial basis neural network, the shape features are used as inputs. The error function is 'mean square error (mse)' which is set to 0.15. The spread for radial basis function is 1.0 and the maximum number of neurons allowed to add during training is 300 [10].

Fuzzy classifier

The fuzzy rule based classification is performed by using mean and standard deviation of the data set (training set). The Sugeno model is used to model any inference system in which the output membership functions are either linear or constant; this model is employed because the expected output is the constant membership function of the class number to which the bacilli cell growth phase belongs. The simple Gaussian membership function

is used and set with the linguistic variables of the mean and standard deviation for the geometric features.

The proposed method for the classification of bacterial cells based on their geometric features is given below:

Training phase:
Algorithm 1: Extraction of features for knowledge base
Step 1 : Input bacterial cell image (RGB color training image)
Step 2 : Convert the RGB image into gray scale image
Step 3: Perform preprocessing method and segment the resulting binary image
Step 4: After removing border touching cells, perform labeling the segmented image
Step 5: For each labeled segment, compute geometric shape features x_i^k , i=1,2,..,5, (i.e. circularity, compactness, eccentricity, tortuosity and length-width ratio) for each cell phase k. The k=1,2,3 correspond to normal, grownup and about-to-divide phase, respectively.
Step 6: Repeat steps 1 to 5 for all the training images
Step 7: Compute mean and standard deviation σ_i^k of the sampling distribution of the feature values for each cell phase k and store them as knowledge base.

Classification phase:
Algorithm 2: Classification of bacterial cell growth phases.
Step 1: Input bacterial cell image (RGB color test image)
Step 2 : Convert the RGB image into gray scale image
Step 3: Perform preprocessing method and segment the resulting binary image
Step 4 : After removing border touching cells, perform labeling the segmented image
Step 5: For each labeled segment, compute geometric shape features x_i ,*i*=1,2,..,5, (i.e. circularity, compactness, eccentricity, tortuosity and length-width ratio) and store these features as $x_i^{(test)}$.
Step 6: Apply 3σ rule for classification of the bacterial cells: A segmented region is of cell phase k, if its features $x_i^{(test)}$ lie in the interval $\pm 3\sigma_i^k$, *i*=1,2,..,5. The k=1,2,3 correspond to normal, grownup and about-to-divide phase, respectively.
Step 7: Repeat the steps 5 and 6 for all labeled segments and output the classification of identified cell growth phases.

The above algorithm for classification phase can be modified to apply k-NN classifier, Neural Network classifier and Fuzzy classifier to the feature set in the Step 6 and the classification performance of the different classifiers, namely, 3σ classifier, k-NN classifier, Neural Network classifier and Fuzzy classifiers can be compared. The k-NN classifier with k=1 is the minimum distance classifier.

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

For the purpose of experimentation, 100 color digital bacterial bacilli cell images containing different growth phase of bacterial cells (non-overlapping) namely, normal, grownup and about-to-divide are considered (as described in section 2). The implementation is done on a Pentium P-IV 1.0 GHz machine using MATLAB 7.9. In the training phase, each input color image of bacterial bacilli cell (Figure 1(a)) is converted into gray scale image (Figure 1(b)), and the morphological operations such as erosion, reconstruction and dilation are applied. The resulting image is thresholded to obtain segmented binary image (Figure 1(c)). The segmented image is labeled and for each segmented region (known cells), the geometric features are computed. The Table 1 presents the geometric feature values computed for the segmented cell regions of the image in the Figure 1(d)-(f).

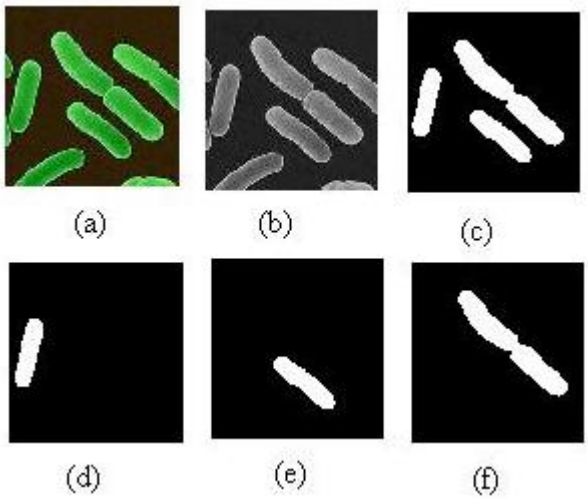
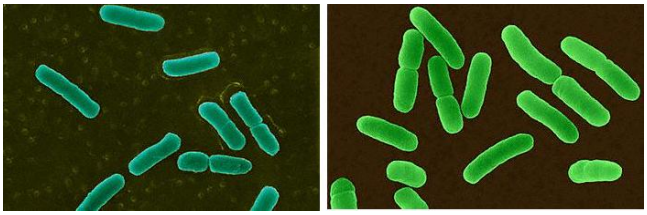


Figure 1. (a) Original color image, (b) gray image of color image in (a), (c) segmented image, (d) the segmented region known as normal, (e) the segmented region known as grownup, (f) the segmented region known as about-to-divide.

The mean and standard deviation of the sampling distribution of these features obtained from the training images are stored in the knowledge base of the bacilli cell growth phases: normal, grownup and about-to-divide, as shown in the Table 2. Some sample training images are shown in the Figure 2.

Table 1. The geometric feature values of the cell regions of the image in Figure 1(d)-(f).

Cell features	Cell growth phases		
	Normal	Grownup	About-to-divide
Circularity (x_1)	0.3227	0.2652	0.1952
Compactness(x_2)	3.0990	3.7710	5.1236
Eccentricity (x_3)	0.9666	0.9792	0.9883
Tortuosity (x_4)	0.3631	0.3735	0.3756
LW ratio (x_5)	3.9014	4.9248	6.5564



In the testing phase, for a test image, the feature extraction algorithm is applied and the test feature values $x_i^{(test)}$ for each segmented region are used for classification using 3σ classifier, k-NN classifier, Neural Network classifier and Fuzzy classifier. The classification results are given in the Table 3 for the testing set images. The Figure 3 shows some sample test images used for classification of bacterial cells.

Figure 2. Sample training images of bacterial cells

Table 2. Mean and standard deviation of geometric features of bacilli bacterial cells with phase types: Normal, Grownup and About-to-divide.

Cell features	Cell growth phases					
	Normal		Grownup		About-to-divide	
	Mean	SD (	Mean	SD (	Mean	SD (
Circularity(x_1)	0.4340	0.0645	0.3411	0.0416	0.2768	0.0388
Compactness(x_2)	2.3516	0.3382	2.9754	0.3801	3.6840	0.5613
Eccentricity (x_3)	0.9385	0.0348	0.9687	0.0080	0.9781	0.0079
Tortuosity (x_4)	0.3713	0.0269	0.3826	0.0254	0.3848	0.0269
LW ratio (x_5)	3.1288	0.6085	4.1274	0.5356	5.0049	0.8414

The Table 3 summarizes the classification accuracy of different classification techniques. The 3σ classifier has yielded an overall accuracy in the range of 93% to 96% and k-NN classifier has yielded 86% to 96% for k=1(i.e. minimum distance classifier). The overall classification accuracy in neural network classifier has yielded 100% in normal cell growth phase, 96% in grownup cell growth phase and 95% in about-to-divide cell growth phase, in fuzzy classifier it has improved and yielded 100% in normal cell growth phase, 98% in grownup and about-to-divide cell growth phases. The performance comparison indicates that the fuzzy classifier has good generalization ability.

Table 3. Classification accuracy for the different bacilli bacterial cell growth phases obtained by different classifiers

Bacterial cell growth phases	No. of cells in test images	Classification accuracy (%)				
		3σ classifier	k-NN classifier		Neural Network classifier	Fuzzy classifier
			k=1	k=3		
Normal	210	96%	96%	92%	100%	100%
Grownup	190	93%	96%	88%	96%	98%
About-to- divide	140	95%	86%	80%	95%	98%

The Figure 4 shows some sample cell images corresponding to misclassification results. In Figure 4(a)-(c), a bacilli cell is not classified (i.e. unknown) due to over segmentation and overlapping. These problems can be overcome by employing better segmentation methods. Further, the classification results can be improved by using better classification techniques. These aspects will be considered in our future work. The Figure 5 shows the top-hat transformed image of the bacilli bacterial cells of Figure 2, which provides the visualization of the different cell growth phases.

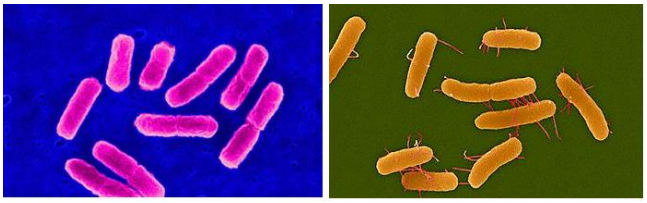


Figure 3. Sample test images used for classification of bacilli bacterial cells

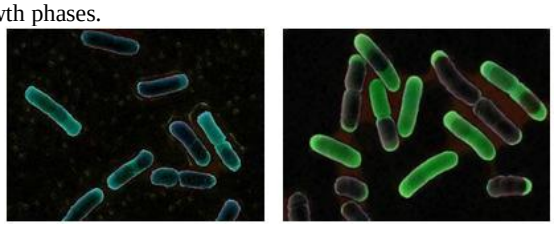


Figure 5 Top-hat transformed image of bacterial cells of Figure 2.

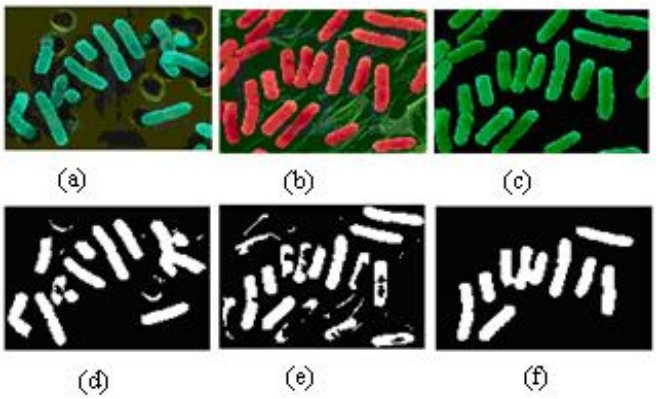


Figure 4. Some sample images corresponding to misclassification results. (a)-(c) original cell images, (d)-(f) segmented images corresponding to cell images in (a)-(c).

5. CONCLUSIONS

In this paper, we have proposed an automated bacilli bacterial cell growth phase classification by segmenting digital microscopic bacterial cell images and extracting geometric features of cells. The experimental results are compared with the manual results obtained by microbiological expert. The proposed method is computationally less expensive and yet yields comparable classification rates in the range 93% to 96% for different cell phases in 3σ classifier, 86% to 96% in k-NN classifier, 95% to 100% in neural network classifier and 98% to 100% in fuzzy classifier. It could be improved further by better preprocessing methods and feature sets, and it could also be employed for different bacterial cell types, which will be taken up in our future work.

6. ACKNOWLEDGEMENTS

The authors are grateful to the referees for their valuable comments and suggestions. Further, the authors are indebted to Dr. A. Dayanand, Professor of Microbiology, Gulbarga University, Gulbarga and Dr. Ramakrishna, Department of Microbiology, Government Degree College, Gulbarga, for providing bacterial bacilli cell images and manual results of the cell images by visual inspection.

7. REFERENCES

[1] Aneja, K. R. (2002). Experiments in Microbiology Plant Pathology Tissue Culture and Mushroom Culture, Newage International Publications, New Delhi, India.

[2] Carolina Wahlby, *et al.*, (2002). "Algorithms for cytoplasm segmentation of fluorescence labeled cells", Analytical Cellular Pathology, **24**, 101-111.

[3] Dennis Kunkel Microscopy, Inc, Science Stock Photography, http://denniskunkel.com/DK/Bacteria/

[4] Hiremath P. S. and Parashuram Bannigidad, (2009). "Automated Gram-staining Characterization of Digital Bacterial Cell Images", Proc. Int'l. Conf. on Signal and Image Processing ICSIP 2009, pp. 209-211.

[5] Hiremath P.S. and Parashuram Bannigidad (2010). "Automatic identification and classification of bacilli bacterial cell growth phases in digital microscopic images", National Seminar on Recent Trends in Image Processing and Pattern Recognition (RTIPPR-2010) , Feb. 15th and 16th, 2010, Bidar, pp.56-59.

[6] Hiremath P.S. and Parashuram Bannigidad (2010). "Automatic identification and classification of Bacterial Cells on Digital Microscopic Images", 2nd International Conference on Digital Image Processing (ICDIP-2010), Proc. of SPIE Vol. 7546-53, Feb. 26-28, 2010, Singapore, pp.754613-1-6.

[7] Liu, J. F.B. Dazzo, O. Glagovela, B. Yu, A.K. Jain (2001). CMEIAS: A Computer-Aided System for the Image Analysis of Bacterial Morphotypes in Microbial Communities, Springer-Verlag, Microb. Ecol. **41**: pp. 173-194.

[8] Madigon M.T. *et al*., Biology of Microorganism, 8th Ed. McGraw Hill Inc., Newyork (1999).

[9] Nicholas Blackburn, *et al.*, (1998). "Rapid Determination of Bacterial Abundance, Biovolume, Morphology, and Growth by Neural Network-Based Image analysis", Applied and Environmental Microbiology, **64(9),** 3246-3255.

[10] Pattan Prakash C., V.D. Mytri and P.S. Hiremath. (2010) "Classification of Cast Iron based Graphite Grain Morpology using Neural Network Approach", 2nd International Conference on Digital Image Processing (ICDIP-2010), Proc. of SPIE Vol. 7546-53, Feb. 26-28, 2010, Singapore.

[11] Petra Perner, (2001). "Classification of HEp-2 Cells using Fluorescent Image Analysis and Data Mining", Medical Data Analysis, Springer Verlag, LNCS **2199**, pp.219-224.

[12] Rafael C. Gonzalez and Richard E. Woods (2002). Digital Image Processing, Pearson Education Asia.

[13] Sigal Trattner and Greenspan (2004). "Automatic Identification of Bacterial Types Using Statistical Imaging methods", IEEETransactions on Medical Imaging", **23(7)**, 807-820.

[14]Thomas Posch *et al.* (2009). "New image analysis tool to study biomass and morphotypes of three major bacterioplankton groups in an alpine lake", Acuatic Microbiol Ecology, **54**: pp. 113-126.

[15]Venkataraman, S., *et al.*, (2006). "Automated image analysis of atomic microscopy images of rotavirus particles", Ultramicroscopy, Elsevier, **106**, 829-837.

[16] S.Osher and J.A.Sethian , "Fronts propagating with curvature dependent speed : Algorithm based on Hamilton Jacobi Formulation" J. Compact Phy, Vol. 79, 1988, pp.12-49.