Software for quantification of labeled bacteria from digital microscope images by automated image analysis

Jyrki Selinummi, Jenni Seppälä, Olli Yli-Harja, and Jaakko A. Puhakka
Tampere University of Technology, Tampere, Finland

BioTechniques 39:859-863 (December 2005)
doi 10.2144/000112018

Automated image analysis software, CellC, was developed and validated for quantification of bacterial cells from digital microscope images. CellC enables automated enumeration of bacterial cells, comparison of total count and specific count images [e.g., 4',6-diamino-2-phenylindole (DAPI) and fluorescence in situ hybridization (FISH) images], and provides quantitative estimates of cell morphology. The software includes an intuitive graphical user interface that enables easy usage as well as sequential analysis of multiple images without user intervention. Validation of enumeration reveals correlation to be better than 0.98 when total bacterial counts by CellC are compared with manual enumeration, with all validated image types. The software is freely available and modifiable: the executable files and MATLAB® source codes can be obtained at www.cs.tut.fi/sgn/csb/cellc.

INTRODUCTION

Direct bacterial counting remains an important tool in microbial ecology. Fluorescent dyes and epifluorescent microscopy are widely used for the enumeration of bacteria. The direct total count (TC) of bacteria can be done by staining microbial nucleic acids with 4',6-diamino-2-phenylindole (DAPI) or acridine orange (1). The fluorescence of DAPI-stained bacteria greatly depends on the sample, and thus the performance of the detection method is important, especially when analyzing bacteria with a low level of fluorescence (2). Fluorescence in situ hybridization (FISH) is used to enumerate specific microbial taxa [specific count (SC)] from natural samples (3). To determine the relative abundance of selected taxa, the same sample is first hybridized with a species-specific probe and then stained with DAPI to account for nonhybridizing bacteria (4,5); this is defined as double staining. After the double staining, the bacteria are enumerated at two different wavelengths at the same position within the sample. The visual enumeration, comparison of

the similarity of TC and SC images, and estimation of cell dimensions are laborious, time-consuming, and difficult tasks. Therefore, automated cytometry is becoming increasingly important.

Digital imaging is one approach to enable the automatic enumeration and analysis of samples. Today, epifluorescent microscopes can be integrated with imaging devices and software for the enumeration of bacteria, including automatic focusing and counting. Previous automated research on microbial enumeration has been primarily based on commercial software (6–11). Microscopic studies are still often done with manually focused microscopes, and commercially available enumeration programs are generally expensive and sometimes complicated to use, especially for a nonprofessional (12). For these reasons, there is a need for free and easy-to-use image-processing software for the enumeration, size measurement, and image-to-image comparison with manual microscopy.

In this paper, we present an easy-to-use cell-counting software tool, CellC. The program includes a

graphical user interface (GUI) and a batch-processing mode that allows for the analysis of multiple images without user intervention. CellC enumerates cells, divides cell clusters into single cells, and compares TC and SC fluorescence microscope images. In addition, quantitative estimates of areas, widths, lengths, and volumes for each cell are produced. The validity of CellC enumeration was assessed by comparison with manual counting results and those acquired by ImageJ (13) image analysis software.

MATERIALS AND METHODS

Method Description

The procedure is divided into two phases: (i) segmentation and analysis of the TC image; and (ii) analysis of the SC image and combining these results with those acquired from the TC image.

The aim of applying image processing to the TC image (Figure 1A) is segmentation of cells from the background and extraction of individual cells from cell clusters to enable data acquisition on a cell-by-cell basis. First, the image background is corrected for brightness variations, appearing in the background because of uneven illumination, and for background autofluorescence. Because the light intensity can be assumed to be quadratic (14), the correction is done by fitting a two-dimensional quadratic polynomial to the image and subtracting the fitted polynomial surface from the original image. Second, the cell pixels are separated from background pixels by global thresholding (15), producing a binarized image with white cells on a black background. Clustered cells are then separated from each other by marker-controlled watershed segmentation (16) that is based on cell intensity. Finally, artifacts such as staining residues are removed by discarding objects smaller than 1/10 of the mean size of all objects or by deleting objects smaller/larger than user-defined thresholds.

When the SC image (Figure 1B) is obtained, the above analysis procedure is repeated on the SC image in a similar

manner. After the analysis, a logical AND operation is performed between the analysis results of TC and SC images. This produces a result with only the objects that were visible in both of the images. The length, width, and volume approximations of the objects are calculated as previously described (17).

Method Validation

For the validation of the enumeration results, two different types of bacterial samples were tested. TC and SC results were validated with bacterial samples taken from a fluidized-bed bioreactor for ferric sulfate production. Samples were visualized by DAPI staining (Figure 1A) and FISH (Figure 1B). For FISH, bacteria in solution were fixed in 4% paraformaldehyde buffered by phosphate-buffered saline (PBS) for 4 h at 4°C before filtering through 0.2-µm pore size polycarbonate filters (Cyclopore® Track Etched Membrane; Whatman, Brentford, UK). The filters were then washed with PBS. The samples were hybridized with CyTM3-labeled oligonucleotide probe for 2 h

and stained with DAPI (1 mg/L for 5 min) (18). In addition, TC results were validated with DAPI-stained filtered bacterial water samples, taken from a pilot-scale water hydraulic system (19). All samples were examined using an Axioskop® 2 epifluorescence microscope, and images were captured with an AxioCam® MPc digital camera using MRGrab 1.0 software (all from Carl Zeiss, Göttingen, Germany). For this study, the bacterial numbers were manually and automatically counted (CellC, ImageJ) from the acquired images.

The Pearson correlation was used as a similarity measure between the results of different enumeration methods. The Pearson correlation coefficient for vectors x and y is defined as

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n - 1)s_x s_y}$$

where n is the number of elements in vectors x , y ; s_x , s_y are the standard deviations of x , y , and $\bar{x}$, $\bar{y}$ are the means of elements of x , y .

RESULTS AND DISCUSSION

Software Description

CellC can be used to analyze bacteria in one image (TC) or to process an image pair (TC and SC). A GUI has been implemented to provide a user-friendly environment for image analysis and to make manual parameter adjustment possible. A batch-processing mode is also implemented, where several images can be processed sequentially without user intervention. The results are saved in a standard comma-separated values (CSV) file that can be opened using spreadsheet software. In addition to the enumeration results, the area, length, width, volume, and intensity measurements for each cell are included in the CSV file. A summary file and binary images (Figure 1, C and D) in JPG format are saved for the validation of the results.

In comparison to the freely available image processing and analysis platform ImageJ (13), our implementation has several differences: CellC has not been designed for basic image processing or for statistical analysis, but rather purely for counting cell numbers and analyzing cell characteristics. This design paradigm enables minimization of different options for the user to define, which makes the software more intuitive for users not familiar with computer science or image-processing terminology. After installation of the software, the analysis procedure is fully automated for color and grayscale images, and the user is automatically prompted to save the resulting analyses. Statistical data analysis is possible by importing the results into spreadsheet software.

In addition to the fully automated analysis, the image-processing procedure can be modified by manually adjusting the analysis parameters. (i) The intensity threshold can be manually selected if the automatic thresholding algorithm fails. The manual threshold adjustment enables the enumeration of cells in cases of a strongly autofluorescent background or in images having low contrast. With the help of a graphical threshold adjuster, changes in the threshold setting can be directly seen. (ii) The automatic background

Figure 1. Examples of original and CellC-rendered result images taken from a fluidized-bed bioreactor. (A) DAPI-stained and (B) FISH images of bacteria taken from the same position (18). (C) Total count (TC) analysis of panel A. (D) Specific count (SC) analysis obtained using Figure 1, A and B. DAPI, 4',6-diamino-2-phenylindole; FISH, fluorescence in situ hybridization.

correction can be set on or off. (iii) The user is given the power to change the degree of division of clustered cells by imposing insignificant minima (16) before watershed segmentation. (iv) The type of microscope image analyzed can be selected: either fluorescence or visible light microscopy. (v) The user can manually set size limits to the enumeration for the cell sizes that are included or the limits can be set automatically. (vi) Finally, the scale can be set as micrometers or pixels. If the scale is not defined, all results are displayed in pixels.

If source code-level analysis of the processing functions is needed (e.g., for details on RGB to grayscale conversion), fully commented MATLAB 7.0 source code has been

released under the GNU General Public License (GPL). This allows users to freely modify the source code and experiment with different algorithms to suit their individual needs. In contrast with the Java language used in ImageJ, MATLAB is a higher level programming language specifically designed for scientific computation. The source code as well as the executable files of CellC are freely available from www.cs.tut.fi/sgn/csb/cellic.

Result Validation

For validation of both TC and SC enumeration results, 19 fluidized-bed bioreactor samples were analyzed. The TC analysis results for comparison of CellC, ImageJ, and manual analysis are presented in Figure 2A. The Pearson correlation, r , in TC analysis between CellC and manual analysis was 0.98; between CellC and ImageJ, 0.99; and between ImageJ and manual analysis, 0.98. The SC results between CellC and manual analysis are illustrated in Figure 2B. For the SC results, the Pearson correlation was 0.92. For additional validation of TC results, 56 images of samples, taken from a water hydraulic system, were enumerated manually and by using CellC. The Pearson correlation between manual and CellC enumeration results was 0.99.

These comparisons show that the results acquired by CellC were very similar to the results of other techniques. Further, because CellC uses a more configurable cell cluster separation algorithm than ImageJ, the procedure can be tuned to produce feasible results from images with small clusters as well as from images with relatively heavy clustering of cells.

The SC images were not processed with ImageJ due to problems in automatic image thresholding (binarization); manual thresholding would have been required to separate the cells from the background. Only

enumeration results were compared between CellC and ImageJ because the results from cell size analysis vary according to the binary thresholding value selected; with high threshold, cell sizes are smaller as more pixels are selected to belong to the background. Small variations in the results with different methods (Figure 2, A and B) were likely due to human error in the manual count and to different image-processing algorithms.

The most significant difference in the image-processing algorithms between CellC and ImageJ is the background removal. For example, in thresholding, because the same threshold value is used in the whole image area, a relatively uniform background is essential for the correct operation of the enumeration and analysis. Part of the difference in the results is also caused by the differences in cluster separation because CellC enables manual adjustment of the cluster separation algorithm.

Future Development

Applications of CellC software are numerous. First, preliminary approximations of cell and object counts can be obtained without expertise in image processing. Second, the verification of previously, manually or automatically acquired results is possible. Third, the summary information produced allows for manual discarding of unreliable results. Fourth, in addition to basic enumeration of objects, the size, intensity, and other detailed information can be obtained. Finally, because no user intervention is needed after the start of the analysis, processing can be done overnight by using a batch of pictures. The correlations of results between CellC, manual enumeration, and ImageJ revealed that CellC produces feasible results.

The main challenge in the analysis and, therefore, the main emphasis of future research, is image segmentation. With strong gradients in the intensity of the image, or strong background autofluorescence that causes poor contrast between the cells and the background, the automated thresholding algorithm is unable to separate the cells from the background. This

Figure 2. Analysis results of 19 bacterial samples taken from a fluidized-bed bioreactor for ferric sulfate production visualized with DAPI staining and FISH. (A) Total cell (TC) counts (DAPI) as measured by CellC, manual analysis, and ImageJ. (B) Specific cell (SC) counts (FISH) from CellC and manual analysis. DAPI, 4',6-diamino-2-phenylindole; FISH, fluorescence in situ hybridization.

also applies to samples with no cells in the image but can be often corrected by manual selection of the threshold value.

Additionally, because cell sizes in microscopic images depend on the staining method used and the size values given by the program are directly measured from the pixel intensities, the staining method may cause underestimation or overestimation of the cell sizes (17). Further work focuses on including additional thresholding and segmentation algorithms in CellC in order to recognize cells from a biofilm, where stained cells have stronger intensity than extracellular polymer substances. In conclusion, CellC enables the enumeration of bacterial cells, provides quantitative estimates of cell dimensions, and includes fully automated analysis as well as manual tuning of the analysis procedure with the help of a GUI. High correlations were observed between the results by CellC, ImageJ, and manual enumeration. CellC can be freely downloaded, modified, and redistributed.

ACKNOWLEDGMENTS

We thank Dr. Sari Soini for her expert advice and Dr. Päivi Kinnunen for providing DAPI and FISH images for the validation of the enumeration results. We also thank Mr. Pekka Ruusuvaara for advice on MATLAB programming. The work was partly supported by the National Technology Agency of Finland (TEKES) and by the Academy of Finland.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

1. Kepner, R.L., Jr. and J.R. Pratt. 1994. Use of fluorochromes for direct enumeration of total bacteria in environmental samples: past and present. *Microbiol. Rev.* 58:603-615.
2. Saby, S., I. Sibille, L. Mathieu, J.L. Paquin, and J.C. Block. 1997. Influence of water chlorination on the counting of bacteria with DAPI (4',6-diamidino-2-phenylindole). *Appl. Environ. Microbiol.* 63:1564-1569.
3. Amann, R.L., L. Krumholz, and D.A. Stahl. 1990. Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. *J. Bacteriol.* 172:762-770.
4. Maruyama, A. and M. Sunamura. 2000. Simultaneous direct counting of total and specific microbial cells in seawater, using a deep-sea microbe as target. *Appl. Environ. Microbiol.* 66:2211-2215.
5. Ishii, K., M. Mußmann, B.J. MacGregor, and R. Amann. 2004. An improved fluorescence in situ hybridization protocol for the identification of bacteria and archaea in marine sediments. *FEMS Microbiol. Ecol.* 50:203-212.
6. Pernthaler, J., A. Pernthaler, and R. Amann. 2003. Automated enumeration of groups of marine picoplankton after fluorescence in situ hybridization. *Appl. Environ. Microbiol.* 69:2631-2637.
7. Gmür, R., B. Guggenheim, E. Giertsens, and T. Thurnheer. 2000. Automated immunofluorescence for enumeration of selected taxa in supragingival dental plaque. *Eur. J. Oral Sci.* 108:393-402.
8. Schönholzer, F., D. Hahn, B. Zarda, and J. Zeyer. 2002. Automated image analysis and in situ hybridization as tools to study bacterial populations in food resources, gut and cast of *Lumbricus terrestris* L. *J. Microbiol. Methods* 48:53-68.
9. Singleton, S., J.G. Cahill, G.K. Watson, C. Allison, D. Cummins, T. Thurnheer, B. Guggenheim, and R. Gmür. 2001. A fully automated microscope bacterial enumeration system for studies of oral microbial ecology. *J. Immunoassay Immunochem.* 22:253-274.
10. Thiel, R. and M. Blaut. 2005. An improved method for the automated enumeration of fluorescently labelled bacteria in human faeces. *J. Microbiol. Methods* 61:369-379.
11. Jansen, G.J., A.C. Wildeboer-Veloo, R.H. Tonk, A.H. Franks, and G.W. Welling. 1999. Development and validation of an automated, microscopy-based method for enumeration of groups of intestinal bacteria. *J. Microbiol. Methods* 37:215-221.
12. Benali, A., I. Leefken, U.T. Eysel, and E. Weiler. 2003. A computerized image analysis system for quantitative analysis of cells in histological brain sections. *J. Neurosci. Methods* 125:33-43.
13. Abramoff, M.D., P.J. Magelhaes, and S.J. Ram. 2004. Image processing with ImageJ. *Biophotonics International* 11:36-42.
14. Klein, A.D., L.R. van den Doel, I.T. Young, S.L. Ellenberger, and L.J. van Vliet. 1998. Quantitative evaluation and comparison of light microscopes. In D.L. Farkas, R.C. Leif, and B.J. Tromberg (Eds.), *Proc. SPIE* 3260:162-173.
15. Otsu, N. 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Systems Man. Cybernet.* 9:62-66.
16. Soille, P. 2003. *Morphological Image Analysis: Principles and Applications*, 2nd ed. Springer-Verlag, New York.
17. Bloem, J., M. Veninga, and J. Shepherd. 1995. Fully automatic determination of soil bacterium numbers, cell volumes, and frequencies of dividing cells by confocal laser scanning microscopy and image analysis. *Appl. Environ. Microbiol.* 61:926-936.
18. Kinnunen, P.H.-M. and J.A. Puhakka. 2004. High-rate ferric sulfate generation by a *Leptospirillum ferriphilum*-dominated biofilm and the role of jarosite in biomass retainment in a fluidized-bed reactor. *Biotechnol. Bioeng.* 85:697-704.
19. Soini, S.M., K.T. Koskinen, M.J. Vilenius, and J.A. Puhakka. 2002. Effects of fluid flow velocity and water quality on planktonic and sessile microbial growth in water hydraulic system. *Water Res.* 36:3812-3820.

Received 4 May 2005; accepted 30 June 2005.

Address correspondence to Jyrki Selinummi, Institute of Signal Processing, Tampere University of Technology, FIN-33101 Tampere, Finland. e-mail: jyrki.selinummi@tut.fi

To purchase reprints
of this article, contact
Reprints@BioTechniques.com