Cell Segmentation for Image Cytometry: Advances, Insufficiencies, and Challenges

Article *in* Cytometry Part A · December 2018

DOI: 10.1002/cyto.a.23686

1 author:

Zhenzhou Wang
Shandong University of Technology
98 PUBLICATIONS **831** CITATIONS

SEE PROFILE

Cytometry PART A
Journal of Quantitative Cell Science

ISAC
International Society for Advancement of Cytometry

Cell Segmentation for Image Cytometry: Advances, Insufficiencies, and Challenges

Zhenzhou Wang*

In the past decades, image cytometry has made great progress due to the advances in optical imaging and image processing. The central problem of image cytometry in many studies is cell segmentation which has received more and more attention in the recent years. The advances of cell segmentation lie in the abundant emergence of new methods, the increased automation level and the increased accuracy. However, most state-of-the-art methods or software tools mainly rely on existing basic image processing algorithms developed many decades ago. For example, the watershed algorithm is adopted by Cellsegm (1), SMASH (2), ImageJ (3), CellProfiler (4), Alanazi's method (5), and the recently reported method by Tsujikawa et al. (in this issue, page XXX). Classical thresholding is adopted by Cellsegm, SMASH, ImageJ, and Alanazi's method while *K*-means clustering is adopted by Tsujikawa's method. These basic image processing algorithms have been developed to solve the general computer vision problems and all of them have inherent insufficiencies. Consequently, they have limited capability to solve the challenging cell segmentation problems at large even if these challenges have been recognized world widely for a long time. The facing challenges include the ever increasing complexity, the great variety of cell types, the low image contrast, the poor image quality, the connection or overlapping of neighboring cells, the nonuniform pixel intensity and the influence of noise or clutter. It is obviously out of the capability of the existing image processing algorithms to solve them at large. Thus, we argue that more effective image processing algorithms should be developed to meet the genericity of the great variety of cells while consistently achieving specificity in solving these challenging problems in different cases.

Overview of State-of-the-Art Cell Segmentation Methods

Cellsegm

Cellsegm (2) could be fully automated or semi-automated based on the user's choice. Its modules include smoothing, ridge enhancement, finding markers, marker controlled watershed segmentation, classification based on user defined threshold, and export of data. After segmentation, the obtained results can be improved by parameter tuning and a recalculation. Cellsegm tends to suffer from serious over-segmentation problems and produce expanded boundaries.

SMASH

SMASH (3) is a semi-automatic MATLAB application that uses histological parameters to segment skeletal muscle. SMASH automates or partially automates determination of cell size, cell type, central nucleated cell, and capillary density. In addition, it allows users to manually control certain aspects of image processing. Its modules include initial segmentation by watershed, cell filtering, cell properties, and segmentation based on Otsu's thresholding algorithm. In its initial segmentation, the watershed may not segment the fiber accurately.

ImageJ

ImageJ (4) is an image processing software tool that makes it possible to solve many image processing and analysis problems. It relies heavily on basic image processing

College of electrical and electronic engineering, Shandong University of Technology, 255000, Zibo City, Shandong Province, China

*Correspondence to: Zhenzhou Wang, College of electrical and electronic engineering, Shandong University of Technology, 255000 Zibo City, Shandong Province, China.
Email: zzwangsia@yahoo.com

Received 23 October 2018; Revised 27 October 2018; Accepted 2 November 2018

Published online in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cyto.a.23686

algorithms. The most frequently used algorithms by ImageJ in cell segmentation are Watershed and Otsu's thresholding. Despite the wide usage of ImageJ, its superiority mainly relies on its comprehensive function instead of the advantages of image processing algorithms.

CellProfiler

CellProfiler (5) is designed for flexible cell image analysis. To effectively identify clumped cells, CellProfiler contains a modular three-step strategy based on basic image processing algorithms, including watershed. CellProfiler tunes the parameters manually by visually inspecting the results. Parameters-tuning was repeated until the best possible segmentation result was inspected. Despite all these efforts, CellProfiler also has over-segmentation problem because of watershed and under-segmentation problem when the cells are connected or overlapped (4).

Alanazi's Method

Alanazi's method (5) consists of max-entropy thresholding and watershed algorithms. High segmentation rates are achieved by employing the optical-phase metric of the transmitted wave-front rather than its intensity. However, their results cannot be immediately generalized to all microbial cell types.

Tsujikawa's Method

Tsujikawa's report presents a new segmentation method for image cytometry analysis (in this issue, page XXX). The parameters are tuned by clustering individual pixels based on the similarity of their useful morphological features extracted with Gabor filters. *K*-means clustering was used to segment the foreground from the background. Aggregated nuclei are discerned using watershed algorithm. This method is affected by the inherent insufficiencies of *k*-means clustering and watershed algorithms.

Wang's Method

Wang' method (6–8) contains three parts and could be divided into three stages. In the first stage, a global threshold is calculated based on the slope difference distribution. For different types of cell images, slope difference-based threshold selection is used to segment the original image and the gradient image, respectively, and the two segmentation results are combined to generate the final segmentation. In the second stage, iterative morphological erosion is used to separate connected or overlapped cells. In the last stage, the boundary of the cell is delineated by contour evolving. Wang's method is capable of segmenting different types of cells robustly (7,8). Figures 1–3 show some typical segmentation and identification results by Wang's method. As can be seen, almost all the cells are segmented successfully and identified correctly. When the differences of the cell sizes in the images vary greatly, over-erosion, or under-erosion problem might occur, that is, the connected or overlapped cells cannot be separated or some cells disappear after the iterative erosion. In such situations, Wang's method could be applied several times with different area thresholds to identify the cells with different sizes, respectively. Despite the relatively high accuracy achieved by Wang's method, it still has a long way to go before this method could achieve high level of automation and high level of accuracy simultaneously for segmenting different types of cells.

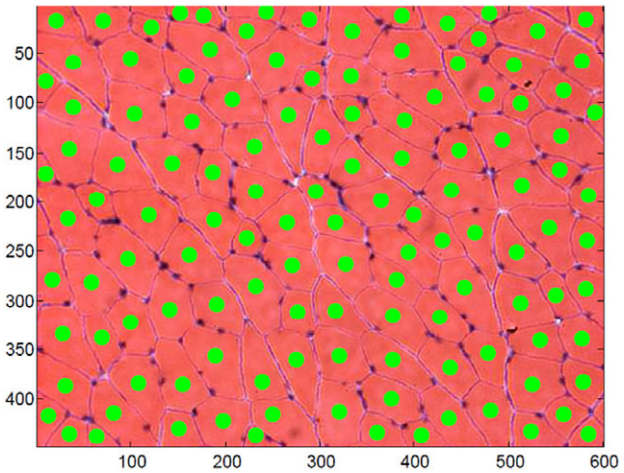

Figure 1. Identification results by Wang's method for muscle cells from Ref. (6).

The Limit of State-of-the-Art Methods

All the referenced state-of-the-art methods, except Wang's method rely heavily on existing basic image processing algorithms proposed in the 1970s for segmentation. All these image processing algorithms have inherent insufficiencies. For example, watershed algorithm is inclined to produce over-segmentation with the increase of the image complexity and the decrease of image quality. On the contrary, the iterative erosion algorithm developed in Wang's method could avoid the over-segmentation problem by selecting proper area threshold to stop the erosion process (6–8). However, fully automatic and unsupervised segmentation by iterative erosion

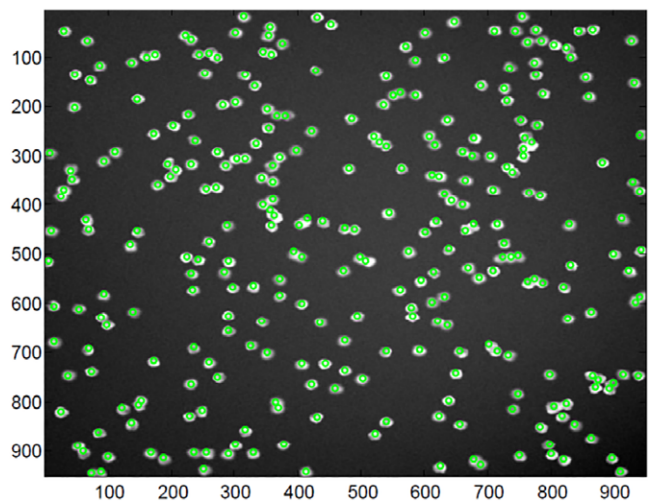

Figure 2. Identification results by Wang's method for synthesized cells from Ref. (9).

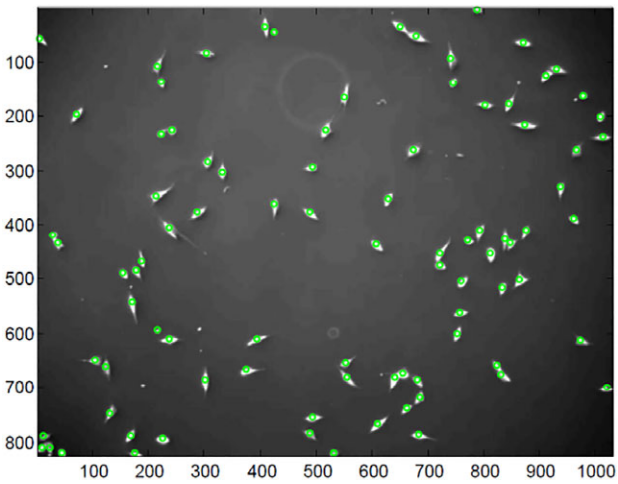

Figure 3. Identification results by Wang's method for pancreatic stem cells on a polystyrene substrate from Ref. (10).

is not achieved yet because of the following challenges. The great variety of cells appears with different shapes, sizes, and characteristics. There might be clutters that are indistinguishable from cells.

Classical thresholding algorithms tend to find the threshold between two pixel classes with larger intensity interval (7). However, the intensity interval between the background and the darker cells might be smaller than the intensity interval between the darker cells and the bright cells in the same cell image. In such situations, classical thresholding algorithms will find the wrong threshold. When the image histogram is not equalized well, classical thresholding algorithms also perform poorly. For instance, Otsu's method could not segment the gradient images with the required accuracy as demonstrated by the experimental comparisons in Refs. (6,7). Furthermore, the accuracy of classical thresholding is easily affected by the overlapping of different pixel classes in the histogram distribution. On the other hand, *K*-means clustering algorithm is also not the ideal option for cell segmentation. *K*-means clustering algorithm will divide the same cell into separate parts when the intensities of different parts of the same cell vary significantly, which is frequently occurring. For the cell image with great noise, both classical thresholding algorithms and *K*-means clustering could not perform well because the noise modifies the histogram significantly and might cause great errors during parameter estimation.

Overall, thresholding is better than clustering for cell segmentation because the gray-scales between different parts of the same cell may vary significantly in some types of cell images. The required thresholding algorithm for cell segmentation should be robust and flexible to adjust to so many types of cell images. Slope difference distribution (SDD)-based threshold selection algorithm developed in Wang's method could meet the requirement of robustness and flexibility (6–8). However, fully automatic and unsupervised segmentation by SDD threshold selection is still difficult to fulfill because of the great variety of cell types, the possible low contrast between the cell and the background, the influence of noise or clutter and the nonuniform pixel intensity in different cells or different parts of the same cell.

The Future of Cell Segmentation

It has taken the community many decades and great endeavor to segment, identify, and analyze cells for explaining the cellular and molecular processes of health and disease. Studies have showed that cell segmentation can contribute to the better cell recognition in various fields of cell biology (1–14) and great advances have been achieved. However, it still has a long way to go before meeting the ultimate goal that the developed method can segment and identify different types of cells autonomously and the biologists can trust the segmentation and identification results blindly. To achieve this ultimate goal, researchers all over the world including biologists, computer scientists, and software engineers should work together closely and industriously.

With the rise of size and complexity of cell images, the requirements for cell segmentation methods are also increasing. The basic image processing algorithms developed decades ago should not be the golden standards for dealing with these challenging cell segmentation problems any more. On the contrary, development of more effective image processing algorithms is more promising for the progress of cell segmentation. In the meantime, comparison of these newly developed algorithms and teaching the biologists to use these newly developed algorithms are also very important. Hence, the open access and authoritative platforms are necessary for researchers all over the world to share, learn, and teach the data, codes, and algorithms.

Literature Cited

1. Hodneland E, Kögel T, Frei DM, Gerdes HH, Lundervold A. CellSegm—A MATLAB toolbox for high-throughput 3D cell segmentation. Source Code Biol Med 2013;8:16.
2. Smith LR, Barton ER. SMASH—Semi-automatic muscle analysis using segmentation of histology: A MATLAB application. Skelet Muscle 2014;4:21.
3. Schneider CA, Rasband WS, Elicieri LW. NIH image to ImageJ: 25 years of image analysis. Nat Methods 2012;9:671–675.
4. Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, et al. CellProfiler: Image analysis software for identifying and quantifying cell phenotypes. Genome Biol 2006;7:R100.
5. Alanazi H, Canul AJ, Garman A, Quimby J, Vasdekis AE. Robust microbial cell segmentation by optical-phase thresholding with minimal processing requirements. Cytometry A 2017;91A:443–449.
6. Wang ZZ. A semi-automatic method for robust and efficient identification of neighboring muscle cells. Pattern Recognit 2016;53:300–312.
7. Wang ZZ. A new approach for segmentation and quantification of cells or nanoparticles. IEEE Trans Ind Inform 2016;12:962–971.
8. Wang ZZ, Li HX. Generalizing cell segmentation and quantification. BMC Bioinformatics 2017;18:189–205.
9. Ljosa V, Sokolnicki KL, Carpenter AE. Annotated high-throughput microscopy image sets for validation. Nature Methods 2012;9:637.
10. Ulman V, Maška M, Magnusson K, Ronneberger O, Haubold C, Harder N, Matula P, Matula P, Svoboda D, Radojevic M, Smal I, Rohr K, Jaldén J, Blau HM, Dzyubachyk O, Lelieveldt B, Xiao P, Li Y, Cho SY, Dufour AC, Olivo-Marin JC, Reyes-Aldasoro CC, Solis-Lemus JA, Bensch R, Brox T, Stegmaier J, Mikut R, Wolf S, Hamprecht FA, Esteves T, Quelhas P, Demirel Ö, Malmström L, Jug F, Tomancak P, Meijering E, Muñoz-Barrutia A, Kozubek M, … Ortiz-de-Solorzano C. An objective comparison of cell-tracking algorithms. Nature methods 2017;14:1141–1152.

11. Allier C, Morel S, Vincent R, Ghenim L, Navarro F, Menneteau M, Bordy T, Hervé L, Cioni O, Gidrol X, et al. Imaging of dense cell cultures by multiwavelength lens-free video microscopy. Cytometry A 2017;91A:433–442.
12. Jørgensen AS, Rasmussen AM, Andersen NKM, Andersen SK, Emborg J, Røge R, Østergaard LR. Using cell nuclei features to detect colon cancer tissue in hematoxylin and eosin stained slides. Cytometry A 2017;91A:785–793.
13. Chiang PJ, Wu SM, Tseng MJ, Huang PJ. Automated bright field segmentation of cells and vacuoles using image processing technique. Cytometry A 2018;93A: 1004–1018.
14. Abas FS, Shana'ah A, Christian B, Hasserjian R, Louissaint A Jr, Pennell M, Sahiner B, Chen W, MKK N, Lozanski G, et al. Computer-assisted quantification of CD3+ T cells in follicular lymphoma. Cytometry A 2017;91A:609–621.