

Understanding the impact of image quality on segmentation accuracy

Ya-Shian Li-Baboud, Antonio Cardone, Joe Chalfoun, Peter Bajcsy, and John Elliott

A method of detecting suboptimal microscopy settings explores the relationship between instrument settings, image quality descriptors, and the accuracy of image post-processing.

High-content screening (HCS) is an automated microscopy technique that enables evaluation of spatial and temporal effects on cells for drug discovery and other applications. The throughput of HCS can be on the order of hundreds of cell images per second to capture transient morphological (structural) effects. Higher throughput can have an impact on image quality by, for instance, reducing exposure times. It can also affect the accuracy of image post-processing, for example, the classification of pixels into biologically meaningful regions, known as segmentation. The motivation of our work is to understand the triangular relationship between image quality, microscopy settings, and accuracy of image post-processing.

Image quality descriptors (such as focus and exposure) augment microscope settings with cell-specific information that not only expands the cell model—i.e., the shape characteristics of a particular line—but is also computationally simple to measure. Previous work has shown how image and region-of-interest (ROI) quality metrics are able to detect suboptimal microscopy settings^{1,2} and how segmentation can impact the accuracy of drug evaluation.³ Figure 1 shows how data quality, a function of both image and ROI quality, can be fed back to the microscopy stage to improve sample preparation and instrument techniques. In addition, data quality descriptors can be used to guide the selection of an optimal segmentation method by assessing the impact of image quality on the accuracy of different approaches.

We have built on previous work by designing a method to detect suboptimal microscopy settings and to derive a computational model based on image quality descriptors to predict segmentation accuracy. Our analysis included two tissue culture cell



Figure 1. Data quality can be related to microscopy instrument settings as well as segmentation accuracy. ROI: Region of interest.

lines (see Figure 2)—A10 and NIH3T3—based on a segmentation study.⁴ Table 1 lists the five microscopy settings we used, which varied the exposure and introduced blurring through a suboptimal filter. Whereas ROI descriptors require detection of foreground edges, image quality descriptors are computed over the entire image. Avoiding the need to detect foreground edges simplifies computation and enables greater flexibility in applications where backgrounds are more complex.

We carried out exploratory data analysis to assess the separability of microscopy settings based on image quality descriptors. Focus (a measure of intensity variance),¹ Blur Power Log,¹ and Entropy⁵ were able to distinguish between the microscopy settings. The Focus descriptor appeared to be a function of exposure settings for both cell lines. The Blur Power Log descriptor proved to be a function of filter type used. The Entropy descriptor is a measure of statistical randomness used to assess the complexity of an image.

Next, we explored the impact of microscopy settings on segmentation accuracy. We use bivariate similarity metrics, TEE and TET, which measure under- and oversegmentation, respectively, to evaluate accuracy.⁴ These metrics provide an understanding of the precision and recall of the number of true pixels in the reference mask compared with the number of estimated pixels

Continued on next page

Table 1. Microscopy settings. A10, 3T3: Tissue culture cell lines. LP: Long pass.

Image condition	Exposure time (A10)	Exposure time (3T3)	Filter type
1	0.015	0.01	Optimal (555nm excitation, 630nm emission)
2	0.08	0.05	Optimal (555nm excitation, 630nm emission)
3	0.3	0.15	Optimal (555nm excitation, 630nm emission)
4	1.0	1.0	Suboptimal (555nm excitation, 740 LP emission)
5	5.0	5.0	Suboptimal (555nm excitation, 740 LP emission)

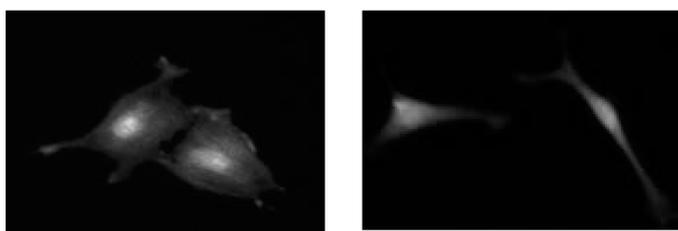


Figure 2. The study was carried out using two types of biological cell lines: A10 (left) and 3T3 (right).

in the segmented mask. The computation of segmentation accuracy results is described elsewhere.⁴

To model the impact of several sources of variation in microscopy settings on segmentation accuracy, we used a multiple regression method. This method examined the weighted combination of image descriptor values and compared it with the accuracy results of each segmentation algorithm tested. A linear combination (i.e., a multivariate weighted model) of image quality descriptors could potentially be used to predict segmentation accuracy for a specific algorithm.

Exploratory data analysis has shown separation of microscopy settings based on box plots of image descriptor values. Single and multivariate analysis also revealed that certain segmentation algorithms are more sensitive to image quality, while algorithms such as the Canny edge detector⁶ are more robust to certain types of image quality degradation. Therefore, we can use the models to determine the best segmentation algorithms given the combination of image quality values.

In summary, the ability to quantify the quality of images leads to better understanding of how image quality degradation due to suboptimal microscopy settings influences the uncertainty of image analysis. Furthermore, determining the impact of image quality on post-processing provides a feedback loop that could improve sample preparation and microscopy settings for future

image sets of cells with similar shape characteristics and thus enhance analysis. Our next steps include evaluating additional image descriptors to broaden our understanding of image quality on segmentation accuracy and to ensure that the method can be scaled to large data sets.

Author Information

Ya-Shian Li-Baboud, Antonio Cardone, Joe Chalfoun, Peter Bajcsy, and John Elliott

National Institute of Standards and Technology (NIST)
Gaithersburg, MD

Ya-Shian Li-Baboud has been a computer scientist in the Information Technology Laboratory at NIST since 2001, serving as principal investigator for a variety of data quality projects in the electronics supply chain, semiconductor manufacturing, and smart grid areas. Her current research interests include data quality, computational biology, computer vision, and machine learning.

Antonio Cardone joined the Information Technology Laboratory at NIST in 2005 and the University of Maryland Institute for Advanced Computer Studies in 2011. His research interests are image segmentation and tracking, computational geometry, and molecular dynamics.

Joe Chalfoun is a research scientist engineer in the Information Technology Laboratory at NIST. He received his PhD in mechanical and robotics engineering and is currently working in the medical robotics field, mainly in cell biology, with a focus on dynamic behavior, microscope automation, real-time tracking, and subcellular feature analysis.

Continued on next page

Peter Bajcsy is a computer scientist in the Information Technology Laboratory at NIST working on automatic transfer of image content to knowledge. His scientific interests include image processing, machine learning, and computer and machine vision.

John Elliott is group leader of the Cell Systems Science Group and a biophysical scientist in the Material Measurement Laboratory at NIST. His research interests include quality control metrics for cell culture.

References

1. M. A. Bray, N. Fraser, T. P. Hasaka, and A. E. Carpenter, *Workflow and metrics for image quality control in large-scale high-content screens*, **J. Biomol. Screen.** **17** (2), pp. 266–274, 2012.
2. A. P. Peskin, A. A. Dima, J. Chalfoun, and J. T. Elliott, *Predicting segmentation accuracy for biological cell images*, **Proc. 6th Int'l Conf. Adv. Vis. Comput.**, pp. 549–560, 2010.
3. A. L. Hill, *Impact of image of segmentation on high-content screening data quality for SK-BR-3 cells*, **BMC Bioinform.**, pp. 340–353, 2007.
4. A. A. Dima, J. T. Elliott, J. J. Filliben, M. Halter, A. Peskin, J. Bernal, M. Kocielek, M. C. Brady, H. C. Tang, and A. L. Plant, *Comparison of segmentation algorithms for fluorescence microscopy images of cells*, **Cytometry A** **79**, pp. 545–559, 2011.
5. Y. Lee, **VASIR: Video-Based Automatic System for Iris Recognition**, PhD thesis, Chung-Ang University, Republic of Korea, 2012.
6. J. Canny, *A computational approach to edge detection*, **IEEE Trans. Pattern Anal. Machine Intell.** **8**, pp. 679–698, 1986.