

## CellAnimation: an open source MATLAB framework for microscopy assays

Walter Georgescu<sup>1,2,3,\*</sup>, John P. Wikswa<sup>1,2,3,4,5</sup> and Vito Quaranta<sup>1,3,6</sup>

<sup>1</sup>Vanderbilt Institute for Integrative Biosystems Research and Education, <sup>2</sup>Department of Biomedical Engineering, Vanderbilt University, <sup>3</sup>Center for Cancer Systems Biology at Vanderbilt, Vanderbilt University Medical Center, Nashville, TN, USA, <sup>4</sup>Department of Molecular Physiology and Biophysics, <sup>5</sup>Department of Physics and Astronomy, Vanderbilt University and <sup>6</sup>Department of Cancer Biology, Vanderbilt University Medical Center, Nashville, TN, USA

Associate Editor: Jonathan Wren

### ABSTRACT

**Motivation:** Advances in microscopy technology have led to the creation of high-throughput microscopes that are capable of generating several hundred gigabytes of images in a few days. Analyzing such wealth of data manually is nearly impossible and requires an automated approach. There are at present a number of open-source and commercial software packages that allow the user to apply algorithms of different degrees of sophistication to the images and extract desired metrics. However, the types of metrics that can be extracted are severely limited by the specific image processing algorithms that the application implements, and by the expertise of the user. In most commercial software, code unavailability prevents implementation by the end user of newly developed algorithms better suited for a particular type of imaging assay. While it is possible to implement new algorithms in open-source software, rewiring an image processing application requires a high degree of expertise. To obviate these limitations, we have developed an open-source high-throughput application that allows implementation of different biological assays such as cell tracking or ancestry recording, through the use of small, relatively simple image processing modules connected into sophisticated imaging pipelines. By connecting modules, non-expert users can apply the particular combination of well-established and novel algorithms developed by us and others that are best suited for each individual assay type. In addition, our data exploration and visualization modules make it easy to discover or select specific cell phenotypes from a heterogeneous population.

**Availability:** CellAnimation is distributed under the Creative Commons Attribution-NonCommercial 3.0 Unported license (<http://creativecommons.org/licenses/by-nc/3.0/>). CellAnimation source code and documentation may be downloaded from [www.vanderbilt.edu/viibre/software/documents/CellAnimation.zip](http://www.vanderbilt.edu/viibre/software/documents/CellAnimation.zip). Sample data are available at [www.vanderbilt.edu/viibre/software/documents/movies.zip](http://www.vanderbilt.edu/viibre/software/documents/movies.zip).

**Contact:** [walter.georgescu@vanderbilt.edu](mailto:walter.georgescu@vanderbilt.edu)

**Supplementary information:** Supplementary data available at *Bioinformatics* online.

Received on June 14, 2011; revised on October 21, 2011; accepted on November 13, 2011

### 1 INTRODUCTION

At present there are a number of microscopy applications on the market, both open-source and commercial, and addressing both very specific and broader microscopy needs. In general, commercial software packages such as Imaris® and MetaMorph® tend to include more complete sets of algorithms, covering different kinds of experimental setups, at the cost of ease of customization and integration with other software. Open-source packages such as CellCognition (Held *et al.*, 2010), CellTracer (Wang *et al.*, 2010), CellTracker (Sacan *et al.*, 2008), Cell-ID (Gordon *et al.*, 2007) or Starrynite (Bao *et al.*, 2006) are easier to customize and integrate well with other software packages for a complete microscopy pipeline. However, these applications are designed for a specific task or set of tasks and cannot easily be made to run other assays. For example, CellCognition, an application designed for high-throughput tracking of fluorescent cells, cannot easily be repurposed to perform a colony counting assay.

Two microscopy applications that offer a high degree of customization are CellProfiler (Carpenter *et al.*, 2006) and ImageJ (Abramoff *et al.*, 2004). CellAnimation is similar to CellProfiler with respect to the modular architecture and high-throughput options. However, CellProfiler is an application whose main purpose is not tracking cells and as such it does not have modules to detect mitotic events, track cells under brightfield illumination or perform tracking correction or virtual staining capabilities. ImageJ offers tracking functionality and other assays through a series of plugins, but is not designed for high-throughput cell tracking or cell ancestry recording.

We have chosen to implement CellAnimation in MATLAB because MATLAB is optimized to work with matrices, it includes a large number of built-in image processing functions and it has been used extensively in the image processing community. This has led to the availability of a large number of thresholding, segmentation and object tracking algorithms. CellAnimation is a framework that provides users with the tools to perform high-throughput cell tracking and ancestry recording, but is suitable for a number of other microscopy assays such as colony and cell counting, focal adhesion tracking, etc. We provide a set of modules (currently >60) and pipelines designed for high-throughput tracking of different types of cells under multiple types of illuminations and the ability to quickly and easily replace parts of these pipelines with higher-performing modules from internal and external sources as they become available. In addition we provide a set of segmentation and tracking correction modules that allow correction and validation of

\*To whom correspondence should be addressed.

automated tracking assays. Finally, our data visualization module allows the selection of multiple cell sub-populations based on shape and motility parameters and the creation of multiple virtual stains that highlight specific sub-populations. We have been using this toolbox in our lab to perform different types of imaging assays both in stand-alone and in a high-throughput mode (Quaranta *et al.*, 2009; Yamashita *et al.*, 2010).

## 2 IMPLEMENTATION

CellAnimation has been developed using MATLAB R2007b and it requires the Image Processing Toolbox. We have run our assays on version R2007b and newer versions (up to R2010b) on Windows, Linux and Mac OS. To allow for fast assay creation and code reuse each assay is implemented as a modular pipeline. A CellAnimation assay is a chain of MATLAB structures. Each structure describes a module and its connectivity. The CellAnimation core functions are responsible for reading the module chain, creating a dependency tree, populating the input values, executing each module and saving those output values required by modules further downstream (Supplementary Fig. 1A). A module is a reusable set of functions that has a narrow specific use (image input–output, thresholding, segmentation, annotation, etc.). In addition we have also included a set of control modules. These are special modules that operate on other modules. Through the use of control modules we can implement looping and branching at the pipeline level. The benefits to this approach are 2-fold. First, the assay logic becomes easier to follow and modify. Second, it allows us to use smaller, more reusable modules.

The modular structure of CellAnimation assays allows for very efficient code reuse. The modules from one assay, such as a cell spreading assay, can be completely reused to create a colony counting assay (Supplementary Fig. 1B). Modifying existing assays or creating new assays from scratch can be easily done using a graphical user interface we provide as part of the framework.

Once an assay has been optimized it can be distributed to a computer cluster for high-throughput processing. Using 60 nodes of the Vanderbilt Advanced Computing Center for Research and Education (ACCRES) cluster, in <2 h we can analyze 60 movies captured over the course of several days and each consisting of hundreds of frames. Since at present it takes about 8 h to transfer the ~100 GB of image data to and from the computing cluster, the speed with which we can analyze movies is limited by network transfer times instead of image processing times!

The performance of assays can be evaluated against manually validated images. To assist with manual validation we have included modules for segmentation and tracking review (Supplementary Fig. 1C and D).

The tracking review module is part of a larger cell-centered image exploration module. The module provides a set of tools that help users explore the patterns and clusters that may be present in the cell population, in addition to correcting errors in automatic tracking. Using the image exploration module a user can explore single cells and population shape and motility parameters or highlight a subpopulation of cells using selection layers (Supplementary Fig. 1C). A selection layer is a transparency overlaid on the original image that highlights certain cells based on a user-defined criterion. The criterion for comparison may be an exact value (all cells with an area larger than 500) or a percentage (cells with an area in the top 20%). Any combination of shape, motility and ancestry parameters may be used alone or in combination to define a layer.

We are in the final stages of integrating our framework with the popular open microscopy environment OMERO to combine high-throughput analysis of datasets stored on an OMERO server, such as happens at Vanderbilt University Medical Center, with the other services provided by the environment, such as dataset annotation.

## ACKNOWLEDGEMENTS

The authors would like to thank Sam Hooke, Shawn Garbett, Yoonseok Kam, Phil Samson and Darren Tyson for their support and suggestions, and Allison Price for her editorial assistance.

*Funding:* National Cancer Institute Grants (5U54CA113007 and 5R01CA047858 to V.Q. in part); Vanderbilt Institute for Integrative Biosystems Research and Education.

*Conflict of Interest:* none declared.

## REFERENCES

- Abramoff, M.D., *et al.* (2004) Image processing with ImageJ. *Biophotonics International*, **11**, 36–42.
- Bao, Z., *et al.* (2006) Automated cell lineage tracing in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. U S A*, **103**, 2707–2712.
- Carpenter, A.E., *et al.* (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.*, **7**, R100.
- Gordon, A., *et al.* (2007) Single-cell quantification of molecules and rates using open-source microscope-based cytometry. *Nat. Methods*, **4**, 175–181.
- Held, M., *et al.* (2010) CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nat. Methods*, **7**, 747–754.
- Quaranta, V., *et al.* (2009) Trait variability of cancer cells quantified by high-content automated microscopy of single cells. *Methods Enzymol.*, **467**, 23–57.
- Sacan, A., *et al.* (2008) CellTrack: an open-source software for cell tracking and motility analysis. *Bioinformatics.*, **24**, 1647–1649.
- Wang, Q., *et al.* (2010) Image segmentation and dynamic lineage analysis in single-cell fluorescence microscopy. *Cytometry A*, **77**, 101–110.
- Yamashita, H., *et al.* (2010) Epitope mapping of function-blocking monoclonal antibody CM6 suggests a “weak” integrin binding site on the laminin-332 LG2 domain. *J. Cell Physiol.*, **223**, 541–548.