
CellTracer: Software for automated image segmentation and lineage mapping for single-cell studies

Quanli Wang¹, Lingchong You² and Mike West¹

¹Department of Statistical Science, Duke University, Durham, NC, 27708, USA.

²Department of Biomedical Engineering, Duke University, Durham, NC, 27708, USA.

July 2008

ABSTRACT

Summary: Imaging of single-cell dynamics is playing an increasingly central role in systems biology studies. Frame-by-frame cell segmentation and lineage analysis from image sequences is central to the goal of data extraction for systems studies, and poses challenges for automating analyses that will apply to multiple studies with differing cell types. From this motivating perspective, we have developed a novel and comprehensive method for automating cell segmentation/lineage reconstruction; this is encoded in CellTracer, a software package to access these methods through an integrated graphical user interface written in Matlab. The software provides manual or semi-automated tools to correct problems from automated algorithms. Studies in dynamic imaging and lineage reconstruction with bacteria, yeast and human cells, as well as multiple other examples of segmentation of different object types, demonstrate the analysis and effective improvements over existing tools as a multi-purpose and easy-to-use analysis approach.

Availability: The software package, manual and examples are available at www.stat.duke.edu/~mw/mwsoftware/CELLTRACER/

1 INTRODUCTION

Cell segmentation and tracking are the two key steps in analysis of dynamic images for data extraction, and are becoming central to single-cell fluorescent microscopy studies in systems biology (Megason *et al*, 2007). Depending on the cells being segmented and tracked, various algorithms and a number of software tools exist in both commercial and public domains (Longo *et al*, 2006). Algorithms often use a combination of morphological operations and some variety of watershed or level set approaches for segmentation coupled with joint estimation of cell correspondences and optimized spatial mappings between frames (Elowitz *et al* 2002). However, algorithm portability and software usability pose major challenges to users due to the different features that these algorithms exploit in their software implementations, the specific customized approaches that often do not easily translate to new contexts, and the requirements for manual, expert oversight and intervention. With these issues in mind and encouraged by the success

in applying our novel, automated hybrid-image based segmentation and neighborhood-score based tracking algorithm to a variety of cell types, we generated CellTracer, a Matlab based graphical user interface program to provide easy access to our automated methodology. With our task-driven modulated algorithm design and interactive user interface, we demonstrate the portability and usability through an array of examples including bacteria, yeast and human cells as well as multiple other examples of effective and, critically, automated image analysis.

2 IMPLEMENTATION AND WORKFLOW

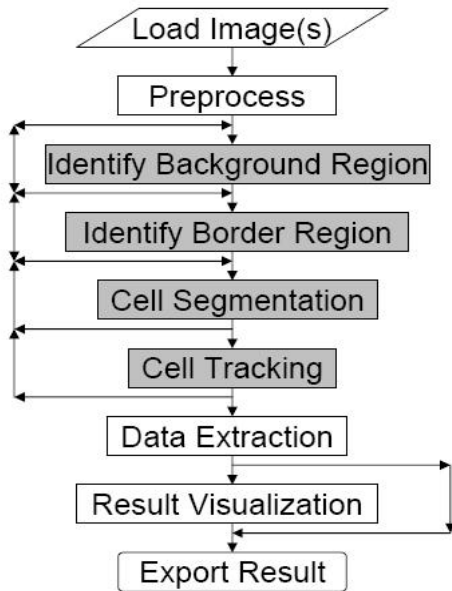
CellTracer is implemented as open source software using Matlab and the Matlab image processing toolbox; it can be run within Matlab or as standalone program. The software features a GUI with standard menus, toolbars and built-in plot panels for rapid visualization, user interaction and result interpretation. It accepts common image file formats as inputs and can export the segmentation/tracking results as images or text files that can be further analyzed by other software. Segmentation and tracking algorithms are packaged as independent, task-oriented tools and can be combined freely to create project-specific workflows to adapt to different input scenarios. CellTracer also enables interactively correction of segmentation/tracking errors throughout the workflow and, if desired (though generally not required) a role for manual control of segmentation and tracking. Detailed descriptions and many examples are available in Wang *et al*, 2008 (submitted).

The CellTracer workflow is depicted in Figure 1A. Input images are preprocessed before segmentation and tracking; this usually involves cropping, inverting and intensity standardization. CellTracer then defines three binary masks to screen the original grey-scale preprocessed images; these represent background, cell borders and cells, respectively. Using the grey-scale images together with their corresponding binary masks, or so-called hybrid images, the automated segmentation method then targets specific image regions -- background, borders and cells -- and iteratively updates the corresponding image masks. This process can be repeated, if necessary, until all the pixels in original images are marked as belonging to one of the three masks. Users can visually

^{*}To whom correspondence should be addressed.

check this process image by image at each step, and manually and interactively correct errors before advancing to next step, although experience with many examples indicates that this is generally unnecessary. For any one application, it may not be necessary to go through all the steps in this general workflow to get satisfactory results. At the end of this process, the resulting cell masks give the segmentation results that are inputs for our neighborhood score-based cell tracking/lineage mapping analysis.

A



B

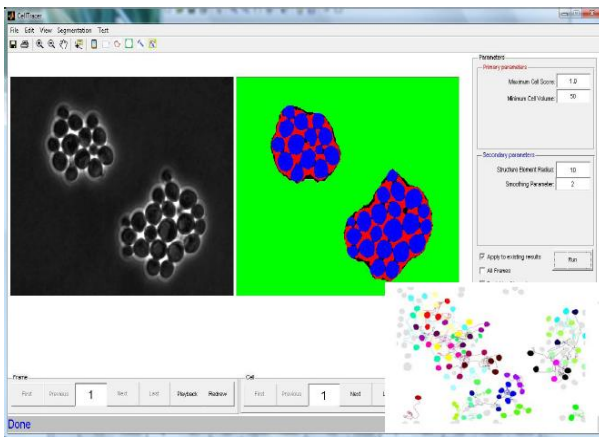


Fig. 1. (A) CellTracer workflow. Grey regions indicate steps that can be manually and interactively intervened upon, if desired. **(B)** Software screenshot, with plot panels showing input image and segmentation result and cell lineage summary plot.

The user can now extract single-cell information in individual images as well as cell-lineage information across image frames and

visualize result in various ways. Figure 1B shows a screenshot of the main CellTracer GUI, with the left panel showing one preprocessed input image and the right panel showing the corresponding segmentation result. An example cell lineage summary plot is also showing in the bottom-right corner of the screenshot.

3 EXAMPLES

CellTracer may be used for a variety of cell types. We have included in the software package three datasets demonstrating its application on *E. coli* bacteria (Rosenfeld, N. *et al*, 2006), budding yeast cells (Di Talia, S. *et al*, 2007) and human cells (Sigal, A. *et al*, 2006) The program gives both segmentation and tracking results from fully automated analyses for these datasets. From analyses with movies of human, bacterial and yeast cells, and many other segmentation studies with other cell types (all exemplified at the CellTracer web site) it is clear that, despite the major differences arising from different cell types, CellTracer is capable of dealing with them satisfactorily using one basic, automated workflow, thus proving its good portability. These diverse examples, drawn from a variety of application areas, can be found at the CellTracer web site along with the software.

FUNDING

CellTracer was partially supported by grants from the National Science Foundation (DMS-0342172, BES-0625213) and the National Institutes of Health (NCI-CA-112952, P50-GM081883-01). Any opinions, findings and conclusions or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of the NSF or NIH.

ACKNOWLEDGEMENTS

We are grateful to J. Niemi and C.M. Tan for useful and formative discussions, to M. Elowitz and N. Rosenfeld for discussion and provision of *E. coli* datasets, and to E. Siggia and S. Di Talia for discussion and provision of yeast datasets.

REFERENCES

- Di Talia, S., Skotheim, J.M., Bean, J.M., Siggia, E.D. and Cross, F.R. (2007) The effects of molecular noise and size control on variability in budding yeast cell cycle. *Nature* **488**, 947-951.
- Elowitz, M.B., Levine, A.J. Siggia, E.D. and Swain, P.S. (2002) Stochastic gene expression in a single cell. *Science* **297**, 1183-1186.
- Longo, D. and Hasty, J. (2006) Dynamics of single-cell gene expression. *Molecular Systems Biology*. **2:64**.
- Megason, S.G. and Fraser, S.E.(2007) Imaging in systems biology. *Cell* **130**, 784-795.
- Rosenfeld, N., Perkins, T.J. Alon, U. Elowitz, M.B. and Swain, P.S. *et al* (2006) A fluctuation method to quantify *in vivo* fluorescence data. *Biophysical Journal* **91**, 759-766.
- Sigal, A. *et al.* (2006) Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins. *Nature Methods* **3**, 525-531.
- Wang, Q., Niemi, J., Tan, C.M., You, L. and West, M. (2010) Image segmentation and dynamic lineage analysis in single-cell fluorescent microscopy. *Cytometry A* **77**, 101-110.