Automated Tracking of Yeast Cell Lineages

Kyungnam Kim*[a], Amy C. Rowat[b], Anne E. Carpenter[c]
[a]HRL Laboratories LLC, Malibu, CA, 90265, USA;
[b]Department of Physics/SEAS, Harvard University, Cambridge, MA 02138, USA;
[c]Broad Institute of MIT and Harvard, Cambridge MA 02142, USA

*kkim@hrl.com; phone 1 310-317-5024; fax 1 310-317-5958; This research work was done while the first author worked at Broad Institute of MIT and Harvard and.

ABSTRACT

We propose a cell progeny tracking method that sequentially employs image alignment, chamber cropping, cell segmentation, per-cell feature measurement, and progeny (lineage) tracking modules. It enables biologists to keep track of phenotypic patterns not only over time but also over multiple generations. Yeast cells encapsulated in chambers of a polydimethylsiloxane (PDMS) microfluidic device were imaged over time to monitor changes in fluorescence levels. We implemented our method in an automated cell image analysis tool, CellProfiler, and performed initial testing. Once refined and validated, the approach could be adapted/used in other cell segmentation and progeny tracking experiments.

Keywords: bio-imaging, time-lapse imagery, cell segmentation, cell tracking, lineage tracking, progeny chambers, yeast.

1. INTRODUCTION

Yeast cell tracking enables biologists to keep track of gene expression patterns of interest in individual cells not only 'over time' but also 'over generations'. We encapsulate yeast cells in chambers of polydimethylsiloxane (PDMS) microfluidic devices, and monitor changes in protein expression after exposure to environmental stimuli or DNA damaging agents. The microfluidic chambers enable maintenance of individual cells in a single focal plane and imaging over multiple generations. We observe how a population of single cells responds to environmental stimuli and the extent to which gene expression patterns are passed on to progeny. Specifically, we measure per-cell fluorescence intensities of specific GFP-fusion proteins (The fluorescence intensities are correlated with protein levels) over multiple generations by tracking cell division in a time-lapse sequence and generating a lineage map that illustrates the growth from ancestor to offspring. One of the key biological questions we try to answer is the following: after exposure to environmental changes such as heat stress, do daughter cells show a similar time-course pattern of protein expression as their mother cell? To what extent is a mother's phenotype inherited by its offspring?

Automated bio-image analysis tools are designed to enable biologists without training in computer vision or programming to quantitatively measure phenotypes from thousands of images. Many bio-image analysis algorithms for time-lapse imagery were reviewed in [1]. Especially, automated cell or cell lineage tracking for biological analysis was studied in [2] (deciphering developmental genes and pathways in *C. elegans* using 3D time-lapse imaging), [3] (cell cycle-dependent changes by measuring protein dynamics in living cells), [4] (quantitative physiology of cellular systems), and [8] (protein localization in budding yeast). General biological object tracking methods were proposed in [5] (Gaussian mixture model for tracking elliptical living objects), [9] (protein movement tracking techniques including optical flow estimation), [10] (4D imaging for precise quantitative analysis). Some cell segmentation techniques were presented in [6] (k-means segmentation on 3D confocal microscopy images), [7] (morphological segmentation), [11] (graphical model based approach), [14] (Voronoi-based technique on image manifolds).

Our proposed yeast cell tracking algorithm consists of the several steps (Figure 1), which we briefly overview here: 1) image alignment; 2) chamber cropping; 3) cell segmentation; 4) per-cell feature measurement; 5) progeny tracking. Step 1 is necessary because a single microscope camera with an automated stage iteratively moves over multiple cell chambers to take time-lapse images at a certain time interval (for example, 10 minutes). Typically, when the stage is re-positioned for the next frame for a chamber, the chamber is in a slightly different location relative to the camera, on the micron-scale. Therefore, the acquired time-lapse images of a particular chamber sequence need to be aligned over time to correctly track those growing cells. In Step 2, the chamber apparatus itself is identified and cropped so that cell

Applications of Digital Image Processing XXXIII, edited by Andrew G. Tescher,
Proc. of SPIE Vol. 7798, 779823 · © 2010 SPIE · CCC code: 0277-786X/10/$18
doi: 10.1117/12.859770

segmentation can occur within the chamber only. For Step 3, segmentation is performed based on cells' intensity and shape using the methods of thresholding and watershed. In Step 4, we measure several features (area, intensity, etc.) for each identified cell. Finally, in Step 5, we use a simple location-based tracking method to identify cells over time and assign new labels to newly budded cells. All the steps are conveniently performed by biologists, using an automated cell image analysis tool, CellProfiler [15] (downloadable at http://www.cellprofiler.org). Even though the algorithm presented here is customized to the yeast cell dataset we collected and works well on those specific images, it is possible to use in automatic progeny tracking of suspension cells (e.g. blood or stem cells) and/or cells in microfluidic channels. Currently, the implemented modules in CellProfiler do not provide a complete solution. Rather, they work as components in a pipeline. There are two independent pipelines – one with the steps from image alignment to feature measurement; the other with progeny tracking. The segmentation and tracking modules are still under development for improvement. Some of the modules can be modified/used in other cell segmentation and progeny tracking experiments as a staring pipeline.

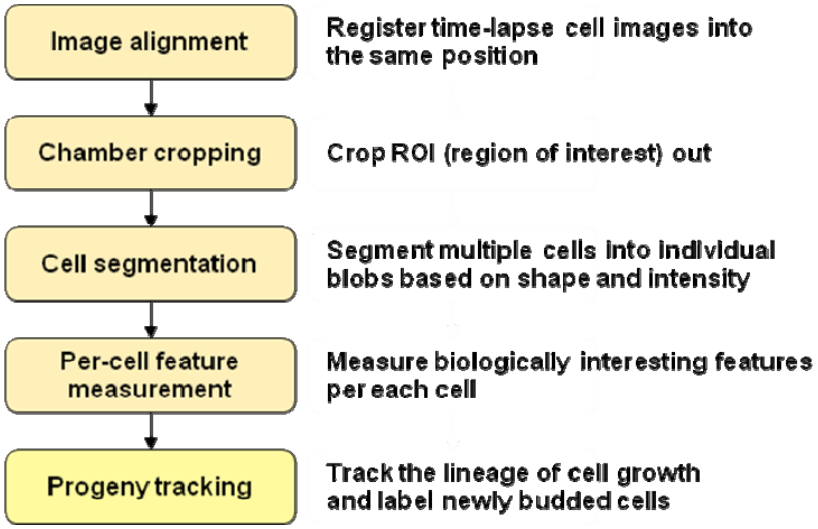

Figure 1 Functional block diagram of yeast cell progeny tracking.

2. IMAGE ALIGNMENT AND CHAMBER CROPPING

Time-lapse images of cell growth (as shown in Figure 2) are captured using an automated robotic stage. When the cells and chambers are imaged by a microscopic camera, they are not perfectly aligned at a fixed position at the micro scale. The images need to be aligned (or registered) into a global coordinate so that cells can be correctly tracked and per-cell feature measurements are accurate.

Because cells are growing and moving over time, those cells themselves cannot be used for alignment. By contrast, the chamber markers as shown in Figure 2 are fixed and clearly visible in the image; these are used for alignment. The most significant transformation factor among the cell images is 'translation' in the x and y directions. The translational offsets Δx and Δy for each image are measured against the reference image, typically the first image of empty chambers. All the input images are aligned to the reference image based on the measured offsets (by applying $-\Delta x$ and $-\Delta y$). The offsets are obtained using the normalized cross-correlation (NCC) method widely used for template matching within a local search window. This algorithm is available within CellProfiler's *Align* module.

Once the alignment step is done, a region-of-interest (ROI) that only contains growing cells in a single chamber is cropped manually out for the subsequent analysis (See Figure 2).

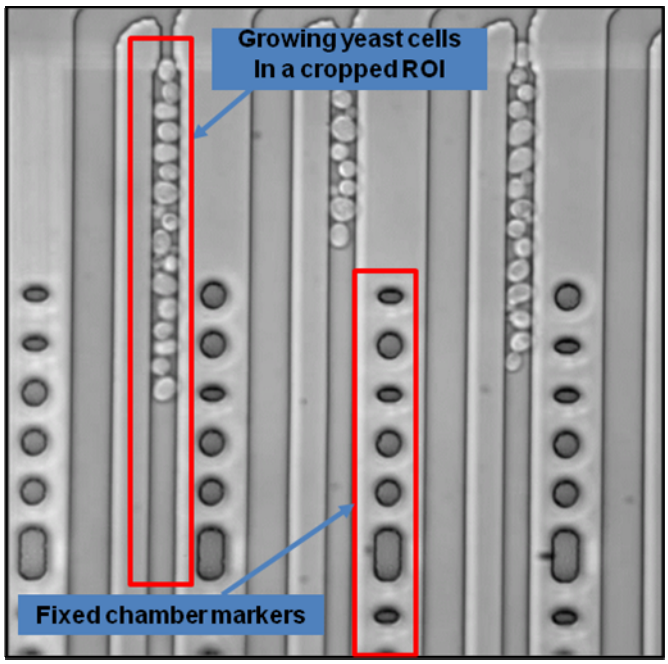

Figure 2 Typical image of yeast cell chambers in a microfluidic device at a single time point. Encapsulated yeast cells in chambers of polydimethylsiloxane (PDMS) microfluidic devices were imaged over time to monitor changes in gene expression.

3. CELL SEGMENTATION

Yeast cells in a cropped ROI image need to be segmented into individual blobs so that per-cell feature measurement can be performed. When capturing time-lapse images of the cell chambers in the microfluidic device, both brightfield and fluorescent (GFP) images are captured from the same cell sample. Green fluorescence protein (GFP) has been widely applied in cell based assays. The fluorescence signals are used as readout to monitor gene functions in a wide variety of research areas such as cell proliferation, differentiation, toxicity, motility, and morphology. We use brightfield images for cell segmentation and GFP images for feature measurement to quantitatively analyze fluorescent cells (see Figure 3). In these experiments, GFP is labeling the hexokinase protein, Hxk1-GFP, that is involved in glucose metabolism and exhibits increased levels during environmental stress response [12].

There are some challenges in segmenting cells in the brightfield images - some edges between cells are fuzzy, cell interiors are not always homogeneous, which creates create false boundaries, cell objects are not much brighter than the background, and cell boundaries are sometimes not clear.

The *IdentifyPrimaryObjects* module in CellProfiler is used for cell segmentation. This module identifies primary objects in grayscale images that show bright objects on a dark background, which is the case in our dataset. Applying automatic thresholding is fast, but fails for these images because cells are touching each other within the chambers. We therefore use the modular three-step strategy in the *IdentifyPrimaryObjects* module. In the first step, the algorithm first determines whether a detected blob is an individual object or two or more clumped objects. This determination can be accomplished in two ways, depending on the cell type: When cells are bright in the middle and dimmer towards the edges, identifying local maxima in the smoothed intensity image works well (*Intensity* option). When cells are quite round, identifying local maxima in the distance-transformed thresholded image (where each pixel gets a value equal to the distance to the nearest pixel below a certain threshold) works well (*Shape* option). In the case of these yeast cell images, the *Shape* option works better to define one marker point per cell.

Then, the edges of cells are identified. For cells within the image that were determined as not clumped, the edges are easily determined using thresholding. For cells that are clumped, there are two options for finding the dividing lines between clumped cells. Where the dividing lines tend to be dimmer than the remainder of the cell (the most common

case), the *Intensity* option works best (the already identified markers are starting points for a watershed algorithm applied to the original image). When no dim dividing lines exist, the *Distance* option places the dividing line at a point between the two cells determined by their shape (the distance-transformed thresholded image is used for the watershed algorithm). In other words, the dividing line is usually placed where indentations occur along the edge of the clumped cells. Both the *Intensity* and *Distance* options work well on the collected yeast cell images.

Lastly, some identified cells are discarded or merged together. Incomplete cells touching the border of the image can be discarded. Objects smaller than a user-specified size range can be discarded. Alternately, any of these small objects that touch valid cells can be merged together based on a set of heuristic rules; for example similarity in intensity and statistics of the two objects. A separate module, *FilterObjects*, further refines the identified cells, if desired, by excluding objects that are a particular size, shape, intensity, or texture. This can be useful to exclude any remaining artifacts.

4. FEATURE MEASUREMENT

Given an image with objects identified, we extract intensity features for each object based on the corresponding fluorescence image. The outlines of cells for measuring per-cell fluorescence intensities are overlaid on the fluorescence image (as shown in Figure 3) for the biologist's information. *IntegratedIntensity* (the sum of the pixel intensities within an object) and *MeanIntensity* (the average pixel intensity within an object) are measured for each cell.

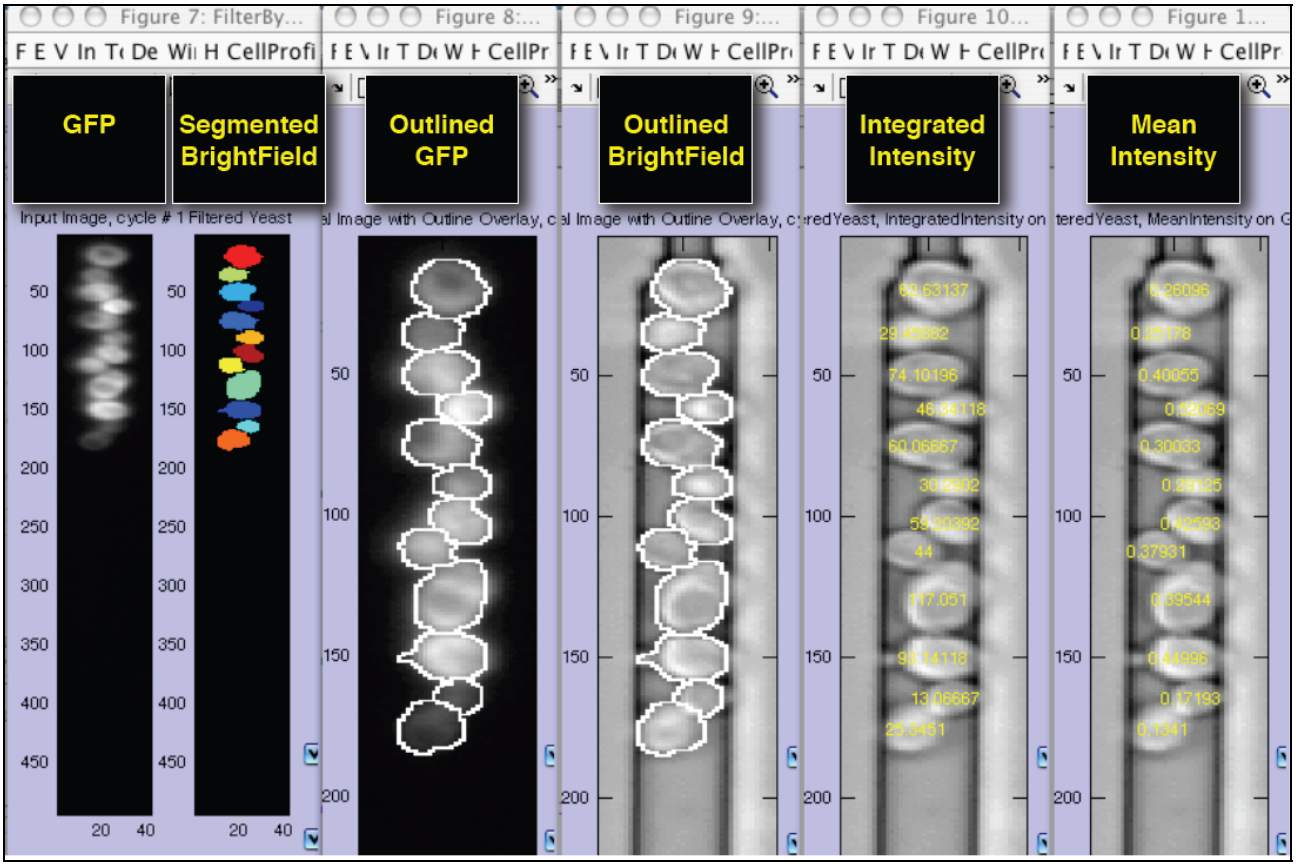

Figure 3 Cell segmentation and feature measurement - Both brightfield and fluorescent (GFP) images are captured from the same cell sample. Objects are segmented into individual blobs. We use brightfield images for cell segmentation and GFP images for feature measurement to quantitatively analyze fluorescent cells.

[Video 1] http://dx.doi.org/10.1117/12.859770.1

5. PROGENY TRACKING

Throughout sequential frames of a cell growth movie, each object is tracked so that it has a stable numerical label in the output measurements. Tracking is performed using the *TrackObjects* module with the *Distance* option. This distance-based method compares the distance between the centroid of each identified object in the previous frame with the current frame. Closest objects to each other will be assigned the same label.

A daughter cell (a new cell budded from an existing mother cell) is labeled with its mother's label as a prefix. The mother cell is determined by the closet cell around the new cell. The daughter cells from the mother cell '1' are labeled such as '1-1', '1-2', and so forth. Thus, the first number represents the ancestor. If there is a single cell at the first frame, all the labels are begins with '1'. Manual progeny re-labeling can be done optionally. After the tracking module automatically labels all the cells, a user can edit errors in cells' labels by selecting them and inputting correct one. The tracking results are shown in Figure 4.

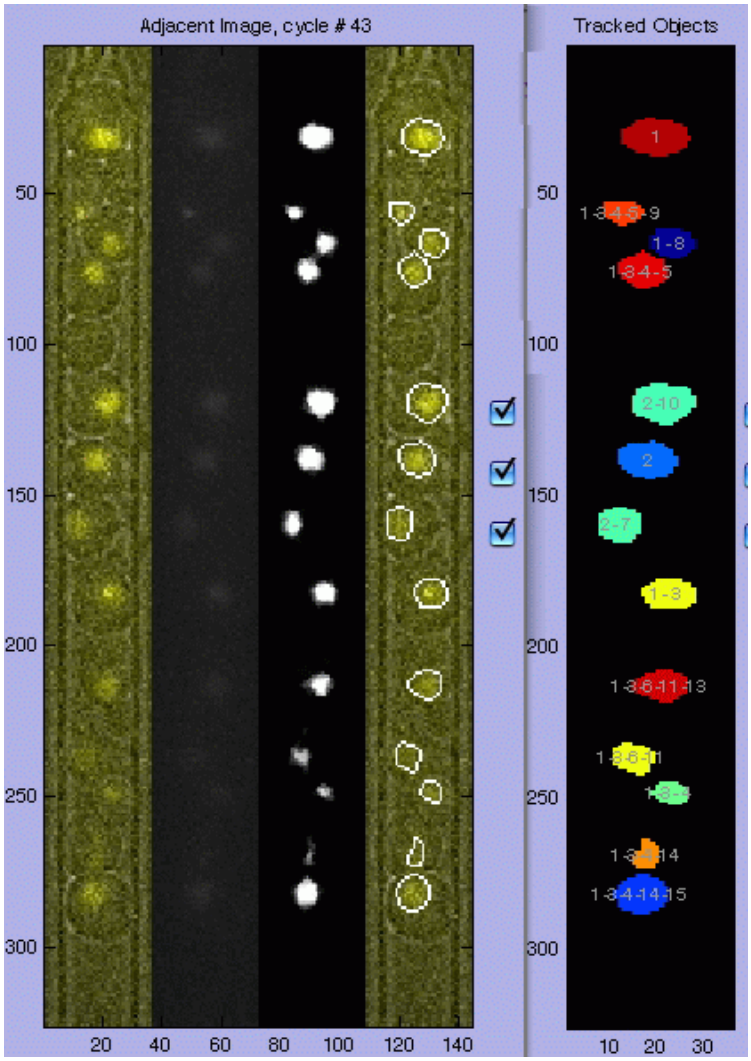

Figure 4 Progeny tracking - each cell is tracked over time throughout out the time-lapse movie and is labeled based on its birth order and mother's label.

[Video 2] http://dx.doi.org/10.1117/12.859770.2

6. CONCLUSION AND DISCUSSION

The proposed yeast cell tracking algorithm with image alignment, chamber cropping, cell segmentation, per-cell feature measurement, and progeny tracking is useful for biologists to keep track of gene expression patterns of interest over time and generations. We observe how a population of single cells responds to environmental stimuli and how gene expression patterns are passed on to progeny using yeast cells in chambers of polydimethylsiloxane (PDMS) microfluidic devices.

The current tracking module uses a naive centroid-based algorithm that looks for the nearest object neighbor from one frame to the next. More robust visual tracking algorithms [13] can handle the cases where there are tracking discontinuities, both spatial (splitting and merging, which requires retaining the history of parents and progeny), and temporal (appearing and disappearing objects; distinguishing situations in which an object has moved outside the field of view or has temporarily dropped to the level of noise, in which case the object trajectory may be reestablished based on object characteristics in neighboring frames). A more advanced tracking algorithm [16] has recently been included in CellProfiler.

REFERENCES

[1] Meijering, E., Smal, I., Dzyubachyk, O., and Olivo-Marin, J.C., "Time-Lapse Imaging" In Microscope Image Processing," F.A.M. Q. Wu, K. R. Castleman, ed. (Burlington, MA: Elsevier Academic Press), pp. 401-440, (2008).

[2] Bao, Z. and Murray, J.I. and Boyle, T. and Ooi, S.L. and Sandel, M.J. and Waterston, R.H., "Automated cell lineage tracing in Caenorhabditis elegans," Proceedings of the National Academy of Sciences, (2006).

[3] Sigal, A. and Milo, R. and Cohen, A. and Geva-Zatorsky, N. and Klein, Y. and Alaluf, I. and Swerdlin, N. and Perzov, N. and Danon, T. and Liron, Y. and others, "Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins," Nature Methods, (2006).

[4] Gordon, A. and Colman-Lerner, A. and Chin, T.E. and Benjamin, K.R. and Yu, R.C. and Brent, R., "Single-cell quantification of molecules and rates using open-source microscope-based cytometry," Nature methods, (2007).

[5] Xiong, G.L. and Feng, C. and Ji, L., "Dynamical Gaussian mixture model for tracking elliptical living objects," Pattern Recognition Letters, (2006).

[6] Niemistö A, Korpelainen T, Saleem R, Yli-Harja O, Aitchison J, Shmulevich I., "A K-means segmentation method for finding 2-D object areas based on 3-D image stacks obtained by confocal microscopy," Proc IEEE Eng Med Biol Soc. 5559-62, (2007).

[7] De Carvalho, M.A.G. and Lotufo, R.A. and Couprie, M., "Morphological segmentation of yeast by image analysis," Image and Vision Computing, (2007).

[8] Chen, S.C. and Zhao, T. and Gordon, G.J. and Murphy, R.F., "Automated image analysis of protein localization in budding yeast," Bioinformatics, (2007).

[9] Miura, K., "Tracking movement in cell biology," Microscopy Techniques, (2005).

[10] Gerlich, D. and Ellenberg, J., "4D imaging to assay complex dynamics in live specimens," Nature Cell Biology, (2001).

[11] Shann-ching Chen , Geoffrey J. Gordon , Robert F. Murphy, "A novel graphical model approach to segmenting cell images," Proc. IEEE Symp. Comput. Intell. Bioinform. Comput. Biol., (2006).

[12] Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO., "Genomic expression programs in the response of yeast cells to environmental changes," Mol Biol Cell.; 11(12):4241-57, (2000).

[13] Alper Yilmaz, Omar Javed and Mubarak Shah, "Object Tracking: A Survey", ACM Journal of Computing Surveys, Dec (2006).

[14] Thouis R. Jones, Anne E. Carpenter, Polina Golland, "Voronoi-Based Segmentation of Cells on Image Manifolds," Computer Vision for Biomedical Image Applications, LNCS Vol. 3765, (2005).

[15] Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, Golland P, Sabatini DM, "CellProfiler: image analysis software for identifying and quantifying cell phenotypes," Genome Biology, (2006).

[16] Jaqaman K., Loerke D., Mettlen M., Kuwata H., Grinstein S., Schmid S. and Danuser G., "Robust single particle tracking in live cell time-lapse sequences." Nature Methods 5: 695 – 702. (2008).