2015

Detecting cells in time varying intensity images in confocal microscopy for gene expression studies in living cells

Debasis Mitra
Florida Institute of Technology

Rostyslav Boutchko
Lawrence Berkeley National Laboratory

Judhajeet Ray
Iowa State University, judhajeet@gmail.com

Marit Nilsen-Hamilton
Iowa State University, marit@iastate.edu

Detecting cells in time varying intensity images in confocal microscopy for gene expression studies in living cells

Abstract
In this work we present a time-lapsed confocal microscopy image analysis technique for an automated gene expression study of multiple single living cells. Fluorescence Resonance Energy Transfer (FRET) is a technology by which molecule-to-molecule interactions are visualized. We analyzed a dynamic series of ~10^2 images obtained using confocal microscopy of fluorescence in yeast cells containing RNA reporters that give a FRET signal when the gene promoter is activated. For each time frame, separate images are available for three spectral channels and the integrated intensity snapshot of the system. A large number of time-lapsed frames must be analyzed to identify each cell individually across time and space, as it is moving in and out of the focal plane of the microscope. This makes it a difficult image processing problem. We have proposed an algorithm here, based on scale-space technique, which solves the problem satisfactorily. The algorithm has multiple directions for even further improvement. The ability to rapidly measure changes in gene expression simultaneously in many cells in a population will open the opportunity for real-time studies of the heterogeneity of genetic response in a living cell population and the interactions between cells that occur in a mixed population, such as the ones found in the organs and tissues of multicellular organisms.

Keywords
Confocal microscopy, FRET imaging, live cell gene expression study, Quantitative biology, Cell detection on time-lapsed image sequence, Time series estimation from image, Scale-space algorithm

Comments
This is a proceeding from *Medical Imaging 2015: Digital Pathology* 9420 (2015): 1, doi:10.1117/12.2081691. Posted with permission.

Detecting Cells in Time Varying Intensity Images in Confocal Microscopy For Gene Expression Studies in Living Cells

Debasis Mitra[1], Rostyslav Bouthcko[2], Judhajeet Ray[3], and Marit Nilsen-Hamilton[3]

[1] Department of Computer Sciences Florida Institute of Technology Melbourne, Florida

[2] Lawrence Berkeley National Lab 1 Cyclotron Drive, Mail Stop 55 Berkeley, California

[3] Ames laboratory and Roy J Carver Dept. of Biochemistry, Biophysics and Molecular Biology Iowa State University, Ames, Iowa

Abstract. In this work we present a time-lapsed confocal microscopy image analysis technique for an automated gene expression study of multiple single living cells. Fluorescence Resonance Energy Transfer (FRET) is a technology by which molecule-to-molecule interactions are visualized. We analyzed a dynamic series of ~10^2 images obtained using confocal microscopy of fluorescence in yeast cells containing RNA reporters that give a FRET signal when the gene promoter is activated. For each time frame, separate images are available for three spectral channels and the integrated intensity snapshot of the system. A large number of time-lapsed frames must be analyzed to identify each cell individually across time and space, as it is moving in and out of the focal plane of the microscope. This makes it a difficult image processing problem. We have proposed an algorithm here, based on scale-space technique, which solves the problem satisfactorily. The algorithm has multiple directions for even further improvement. The ability to rapidly measure changes in gene expression simultaneously in many cells in a population will open the opportunity for real-time studies of the heterogeneity of genetic response in a living cell population and the interactions between cells that occur in a mixed population, such as the ones found in the organs and tissues of multicellular organisms.

Key Words: Confocal microscopy, FRET imaging, live cell gene expression study, Quantitative biology, Cell detection on time-lapsed image sequence, Time series estimation from image, Scale-space algorithm

1. Description of Purpose

Fluorescence Resonance Energy Transfer (FRET) is a technology by which molecule tomolecule interactions are visualized [Förster'48]. Confocal microscopy enables the study of gene expression in single living cells utilizing FRET emission. The technique [Shin'14] involves the production of aptamers as a result of gene expression, which bring together two dyes (Cy3 and Cy5) that interact in FRET. The cells expressing the aptamers are illuminated to activate the donor dye (Cy3), which then interacts with the acceptor dye (Cy5) that releases the energy at an emission wavelength characteristic of the acceptor. The release of energy to the acceptor results in less fluorescence emission by the donor. Thus, to measure FRET, images are simultaneously collected of the intensity of emission of the donor and the acceptor (see Fig. 1). Image analysis can involve a significant amount of manual processing that may take days for each experiment. To provide the flavor of the difficulty level of a sample task: there are three image sequences, each sequence consists of sixty-three microscopic images of hundred-plus cells, and each cell may span over tens of pixels. Pixels' values over a cell need to be collected (and averaged over the cell's spatial extent) to quantify the concentration of the relevant fluorescing molecules in the cell. Moreover, the cells spontaneously move horizontally in the image space. A cell may also move vertically in the gel causing its shape, size, and intensity to vary on the focal plane of the microscope. Cells may overlap with each other in the focal plane, which further complicates the ability to distinguish them. While the manual analysis produces high quality results when performed by an expert, the process takes many days for each study. Our project is to automate this process and we report here an algorithm for that purpose. The main computational challenge is to locate the individual cells over multiple frames, and produce a time series for each image per cell.

2. Methods:

Imaging details: Yeast cells, transformed with an expression vector for 6xPDC IMAGEtags under the control of the GAL1 promoter, were grown for 12h in the presence of 2% raffinose to an OD_{600} of 0.9. Cy3-PDC (10 μM) and Cy5-PDC (4 μM) were added and cells incubated for 90 min before the addition of 2% galactose to activate gene expression. Images were taken starting 6 min after the addition of galactose and were collected at intervals of 1 min or 3 min over most of the 142 min imaging period. Channels for Cy3 excitation/Cy3 emission were 550 nm/560-626 nm; for Cy5 excitation/Cy5 emission were 650 nm/660-754 nm; and for Cy3 excitation/Cy5 emission were 550 nm/660-754 nm.

The four channel image was generated by a Leica SP5X laser scanning confocal microscope with a 63X objective and immersion oil. Sixty three time-lapsed frames of 512x512 pixels from each channel were first extracted as an image-sequence by preprocessing – total four sequences. We call them c0 (Cy3 excitation Cy3 emission), c1 (Cy5 excitation

Medical Imaging 2015: Digital Pathology, edited by Metin N. Gurcan, Anant Madabhushi, Proc. of SPIE Vol. 9420, 942011 · © 2015 SPIE
CCC code: 1605-7422/15/$18 · doi: 10.1117/12.2081691

Cy5 emission), c2 (DIC, differential interference contrast), and c3 (Cy3 excitation Cy5 emission). C3 is used to detect FRET output, and c2 shows all cells in the field and is used to verify that each region of FRET signal is associated with a cell. C2 was not used for the quantitative analysis. Our task is to estimate the average intensity values of each cell across time and across the three channels, c0, c1, and c3.

In order to eliminate the background, the sequence c3 is first thresholded with a cutoff value carefully chosen by observation of the histogram of the intensity values. Next a *scale-space* blob detection algorithm [Lowe'04] is used to find the cells on each 2D frame independently. Scale space algorithm uses multiple varying width Gaussian windows to convolve over each 2D frame and chooses the cell location by the largest value from convolution. This algorithm finds the center and radius of each cell as a circle. In future we intend to extend the algorithm to detect cells as ellipses, which is more suitable for their shapes in the images. After this stage, some cells may be found overlapping with each other on the frame. We remove the cells that overlap with more than 30% area with another cell, by choosing only the larger one of them. However, until this point the cells are chosen based only on each 2D frame independently. They need to be subsequently 'threaded' on time-lapsed frames, and as they may move in 3D space this is not an easy task. To recognize each cell across the frames we first converge each cell's average intensity value on its center pixel in each frame and put zero values everywhere else. Thus, we create an average intensity-image sequence with zero values everywhere other than on the center of each cell, which is a single pixel with its average intensity value. We presume each non-zero value on the first frame potentially as a cell. Next, we use a threading-like operation across the channels looking for the non-zero value within a pre-assigned radius, going from each frame to the next frame only. The real challenge is when more than one non-zero pixels are encountered within the sought-after radius in the next frame. A simple heuristic of using the closest value (average intensity) to the previous frame is chosen as the tie breaker for identifying the center of the same cell in the next frame. Since a cell may also disappear over one or a few frames, we need to look for its presence over not only on the next frame but also over a certain number of subsequent frames, and if not found by then, we give up on that cell. Also, if a cell does not show up on more than of a total 40 out of 63 frames, then that cell is discarded from our consideration. At the moment, we do not care about any cell that does not appear on the first frame, which is weakness of the process. In future we will fix this. All the above procedure is performed on the image sequence c3. Once the cells' centers and radii are detected across the frames on c3, we use these positions to collect average intensity values on the other two channels' sequences c0 and c1, and report them on a spreadsheet for further quantitative analysis.

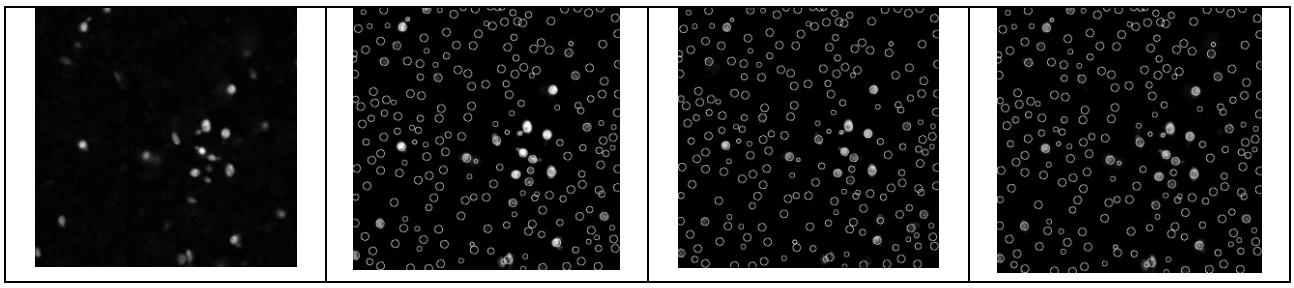

Figure 1 (from left upper). Image c3 (Cy3 excitation Cy5 emission); Cells on c3 (Cells are identified and the estimated circular extents, as obtained by the scale-space algorithm, are highlighted with a border); Cells on c0 (Cy3 excitation Cy3 emission, same cell locations as obtained from c3); Cells on c1 (Cy5 excitation Cy5 emission, same cell locations as obtained from c3).

3. Results:

Fig. 1 shows a sample frame of c3 sequence (left) and the scale-space algorithm recognized circular spans of the cells on it (second). Next two images of Fig. 1 show the same cell locations on the same frame of c0 and c1 sequences.

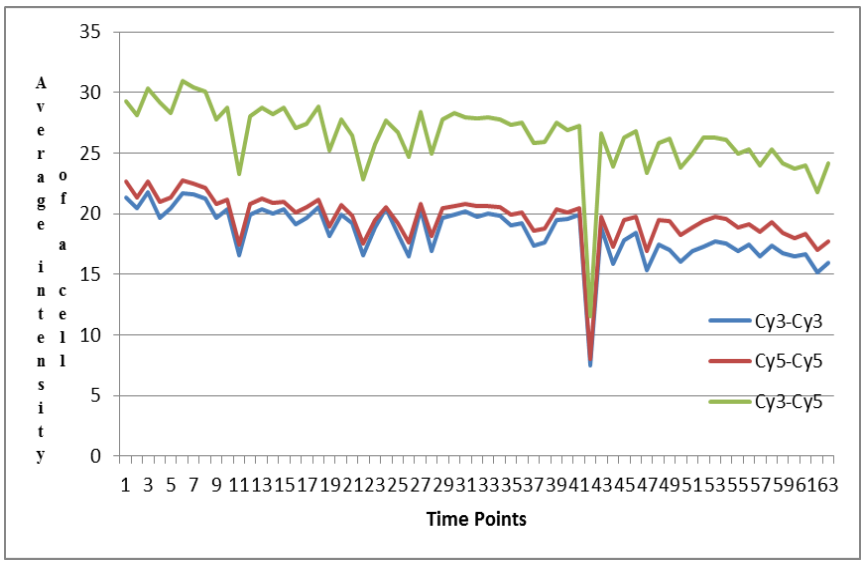

Figure 2: A sample cell's average intensity variation over time-lapsed frames. Three colored curves represent three channels: Cy3 excitation Cy3 emission, Cy5 excitation Cy5 emission, and Cy3 excitation Cy5 emission.

Fig. 2 shows unsmoothed time series for a sample cell over three types of images (channels): Cy3 excitation Cy3 emission, Cy5 excitation Cy5 emission, and Cy3 excitation Cy5 emission. The cell's vertical movements across the focal plane are quite apparent from the fluctuation of the corresponding intensity values. Some smoothing process over the time-series may remove the fluctuations, but may affect further quantitative study and may not be necessary.

The algorithm is implemented in Matlab™ and takes about three minutes for the pre-processed (four channels separated) data.

4. Summary of breakthrough

Our main methodological contribution is to replace a laborious, expensive, and possibly subjective data analysis protocol, with a simple efficient algorithmic process. Many complicated algorithms have been developed [Chang'08, Amat'13] for detecting cells on microscopic images, but for our purpose in confocal imaging the algorithm based on the scale-space technique as presented here works quite well. It has many directions for further improvement, e.g. it is a highly parallelizable algorithm, thus with a possibility to improve the speed to the extent that it may be embedded in the firmware of a microscope. In addition to applying for FRET imaging, this algorithm has equal utility for the analysis of cells that are visualized by confocal microscopy such as those illuminated for fluorescence imaging by their expression of one of many available fluorescent proteins. The ability provided by this method to quickly gather data from many cells in the one image sets the stage for analyzing gene expression in a population of living cells in real time as they interact with their environment and with adjacent cells. The tissue microenvironment and the nearby cells can greatly impact the behavior of cells in living tissues and organs of multicellular organisms. This method of tracking gene expression in living cells will provide a means for studying and understanding diverse events in microbes, animals and plants that are influenced by the microenvironment in which the cells exist. Examples of such events include those involving cell interactions during normal tissue development and maintenance, the growth and metastasis of cancers, the interaction of cells in plant root hairs with surrounding microbes, and interactions between microbes in biofilms.

5. Conclusion

In this work we have developed an algorithm and implemented it in Matlab to analyze large quantity of FRET imaging data from confocal microscopy for studying gene expression in living cells.

6. Acknowledgment

Research presented in this manuscript is supported in part by the Director, Office of Science, Office of Biological and Environmental Research of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract No. DE-AC02-07CH11358.

Disclaimer: This work has not been presented to any other forum so far.

References

[Amat'13] Amat F., and Keller PJ. "3D HAAR-like Elliptical Features for Object Classification in Microscopy," *IEEE 10th International Symposium on Biomedical Imaging*, 2013.

[Chang'08] Chang H., Yang Q., and Parvi B. "A Bayesian Approach for Image Segmentation with Shape Priors," *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, pp 1-8, 2008.

[Forrster'48] Förster, T. "Zwischenmolekulare Energiewanderung und Fluoreszenz" [Intermolecular energy migration and fluorescence]. *Annalen der Physik* (in German), 1948.

[Lowe'04] Lowe D. "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, Volume 60(2), pp. 91-110, 2004.

[Piersma'03] Piersma S., Denham E.L., Drulhe S., Tonk R.H.J., Schwikowski B., and, van Dijl J.M. "TLM-Quant: An Open-Source Pipeline for Visualization and Quantification of Gene Expression Heterogeneity in Growing Microbial Cells," *Plos One*, Volume 8, Issue 7, 2003.

[Shin'14] Shin, I., Ray, J., Gupta, V., Ilgu, M., Beasley, J., Bendickson, L., Mehanovic, S., Kraus, G., and Nilsen-Hamilton, M. "Live-cell imaging of Pol II promoter activity to monitor gene expression with RNA IMAGEtag reporters," *Nucleic Acids Res*. Volume 42, Issue 11, e90, 2014.