

Segmentation and Cell Tracking of Breast Cancer Cells

Adele P. Peskin¹, Daniel J. Hoepfner², and Christina H. Stuelten²

¹ NIST, Boulder, CO 80305

² NIH, Bethesda, MD 20892

Abstract. We describe a new technique to automatically segment and track the cell images of a breast cancer cell line in order to study cell migration and metastasis. Within each image observable cell characteristics vary widely, ranging from very bright completely bounded cells to barely visible cells with little to no apparent boundaries. A set of different segmentation algorithms are used in series to segment each cell type. Cell segmentation and cell tracking are done simultaneously, and no user selected parameters are needed. A new method for background subtraction is described and a new method of selective dilation is used to segment the barely visible cells. We show results for initial cell growth.

1 Introduction

The ability of cells to migrate varies widely during development and the life of multicellular organisms. For example, cell migration is crucial during the early stages of development when organs are formed. During adulthood, a wide array of immune cells can migrate at speeds that can reach more than one cell length per minute. These cells travel over wide distances to sites of inflammation or injury to precipitate the host's immune response. On the other hand, epithelial cells, which line cavities and surfaces inside the body, are characterized by tight contact with neighboring cells and typically do not migrate over far distances.

Cell- and organ-specific regulation of the migratory behavior of cells is crucial under both physiological and pathological conditions. In normal epithelia, such as breast epithelium, cell migration is restricted, but in tumor disease these transformed epithelial cells acquire mutations that allow them to migrate into the surrounding connective tissue (invasive tumor) and then to distant sites to establish metastatic disease [1]. Therefore, investigating the migratory behavior of epithelial and tumor cells may enable us to find new ways to control migration of tumor cells to distant sites and reduce metastasis.

Cell migration can be studied by visualizing cell responses through the capture of time-lapse images for long periods of time. Quantitative analysis of these images is performed by tracking the position of individual cells as a function of time, which is labor-intensive and can only be done manually for a few selected

⁰ This contribution of NIST, an agency of the U.S. government, is not subject to copyright.

cells. Thus, current ways to analyze cell migration often generate data of limited statistical value. To robustly analyze migration of cells over a time range of days, an image analysis program would ideally be able (a) to segment images and identify cells, (b) to individually name cells, (c) to assign cells in one image to cells in the preceding and succeeding images, that is, identify the same cell in sequential images although this cell may change location, size and shape, (d) identify whether a cell dies, (e) notice whether a cell divides and identify the offspring as new cells but related to the parental cell.

We developed a segmentation method specific for the cell line studied here, MCF10CA1a cells [2]. Many segmentation methods exist in the literature: edge-based methods, methods which minimize certain energy functions [3] [4], and methods that combine different techniques, such as the coarse-to-fine iterative segmentation of 3T3 cells [5]. Several methods determine cell boundaries by tracking cell movements over time, either by tracking cell trajectories [6], or by feature-specific shape change detection [7]. Ersoy et.al. propose a flux-tensor based method that detects cell movements [8]. The images from this cell line present segmentation problems that are not addressed by these solutions: cell edges are not clearly defined and cell shapes are not consistent from one image to the next, so methods that depend on these features are not useful for this cell line. Other cell tracking methods fail to accurately track both the bright round cells, and the barely visible cells. In this paper, we discuss a method that tracks the combination of diverse, often fast-moving, cells of this breast cancer line. The parameters used in the segmentation method are based on the statistics of image intensities and user input is not required.

2 Tissue Culture

Breast cancer cells (MCF10CA1a cells) were cultured as described previously [9] [10]. Briefly, cells were grown in DMEM/F12 supplemented with 5 % horse serum (both Invitrogen, Carlsbad, CA) in a humidified atmosphere at 5 % CO₂. For experiments, cells were trypsinized and resuspended in a 1:1 mixture of DMEM/F12 supplemented with 5 % horse serum and DMEM, low glucose, supplemented with 10 % fetal bovine serum (Invitrogen). 25000 cells were plated into each dish (total volume: 3 ml). Cells were allowed to adhere overnight and then transferred into the incubator chamber of the microscope for monitoring cell migration.³

3 Microscopy

Cultured cells were imaged continuously in a 5 % CO₂ environment (Zeiss CO₂ module S). Temperature was regulated to 37 °C ± 0.1 °C by use of an AirTherm

³ Certain trade names are identified in this report in order to specify the experimental conditions used in obtaining the reported data. Mention of these products does not constitute an endorsement of them.

heater (WPI). Cells were imaged by use of a Zeiss AxioObserver microscope with a 40x, 1.3 NA oil-immersion objective and Cargille 37 immersion oil. DIC images were captured with an Orca EM (Hamamatsu) CCD camera every 2 min and saved as .zvi files at 512x512 pixels and 16 bit depth. Original files were converted to .tif series using ImageJ with no post-hoc file manipulation. Multiple areas of the coverslip were imaged simultaneously using a computer-controlled motorized stage (Ludl Mac5000) with a linear encoder. Axiovision (Zeiss, Version 4.6) software was used to control all hardware components. To minimize photo toxicity, 15 ms exposure times were achieved using a shutter (UniBlitz). Sample images can be found at: <http://www.nist.gov/itl/math/hpcvg/cellingvis.cfm>.

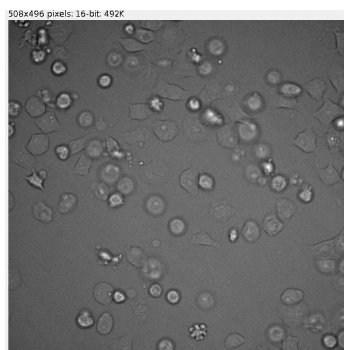


Fig. 1. Early image.

4 Background Removal

Background values for live cell images have been computed in a number of different ways, from taking the mean value or weighted mean of pixel intensities at each location [11] [3], to modeling pixel intensities over time at each location with Gaussian distributions representing cell and background [12], to top-hat filtering [4]. The goal for this segmentation is to do simultaneous cell tracking along with the segmentation, so we chose a method that can be performed on each image without knowledge of previous background image information. Background removal needs to be highly accurate in order to track the barely visible cells. To start, a mean value is found for the image as a whole, and all pixels with an intensity higher than three standard deviations above the mean are temporarily replaced by the mean value to smooth out the image. Then the background value for each individual pixel is determined by the average value in a small neighborhood of that pixel. We experimented with different sized neighborhoods, from 3 to 20 pixels in each direction. A neighborhood of ten pixels in each direction was used, which produced a smooth background without removing features of the image. Figure 1 shows an image at the start of growth

in new cell medium, a 16-bit 508x496 pixel image (intensity units are on a scale of 0-65536). Background removal for this image is displayed in Figure 2. The presence of the barely visible cells in the resulting image indicates the success of this technique.

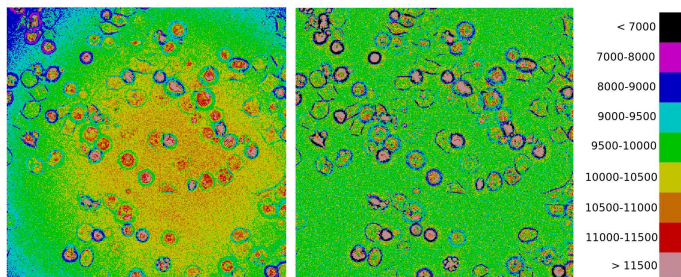


Fig. 2. Sample image color-coded before and after background removal. Color table shows intensities before removal.

5 Initial Segmentation

The initial images in a time sequence contain cells of a fairly similar size, although even at this stage several quite different cell morphologies are seen. Figure 3 shows some closeup pictures of a few of the cells from Figure 1, showing both cells with clearly defined boundaries, and cells that are barely visible on the image with large gaps in any apparent boundary. The overall approach is to capture the dark cell boundaries to define the brighter cells, and then use a combination of selectively dilating the fainter bright areas that mark the barely visible cells with a measure of pixel intensity standard deviation, which also indicates faint areas on the image that represent cells. Dilation in this paper will refer to enlarging a set of selected pixels by including pixels directly above, below, to the left or to the right of a selected pixel. Erosion refers to decreasing a set of selected pixels by eliminating pixels that have an unselected pixel above, below, to the left or to the right.

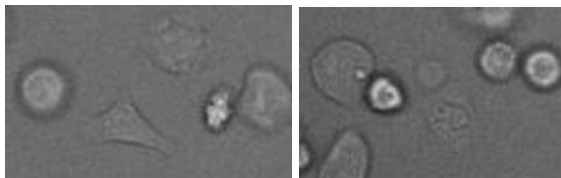


Fig. 3. Sub-sections of the image in Figure 1, showing different cell morphologies.

To begin we define several overall patterns within the image, starting with the areas of high standard deviation. We want to capture cell features on the scale of the size of a dark boundary or a bright shadow inside a dark boundary. Most of these features are represented by 3-10 pixels on the image. We experimented with neighborhoods varying between 3 and 15 pixels in each direction, looking for a pattern that best resembled the visual pattern of the cells. For each pixel, intensities in a neighborhood of five pixels in each direction are collected, and a mean and standard deviation are assigned to that pixel. The individual pixel values are then averaged to give an overall mean (om) and overall standard deviation for the pixel standard deviations (osd). Figure 4 shows pixels with individual standard deviation values above the overall mean plus one overall standard deviation ($om + osd$).

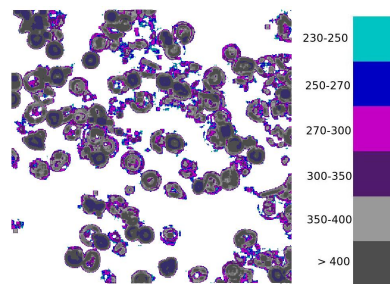


Fig. 4. Pixels with high standard deviation values.

The dark areas defining cell borders are found in a series of steps shown in Figure 5. We look for the lowest possible intensity cutoff that leaves complete cell boundary sections intact and choose all pixels less than -100 intensity units after background removal. We remove small clusters of dark pixels. Most of the dark pixels of the background are present in small clusters of 10-15 pixels or fewer, so all remaining clusters of 15 or less of these pixels are eliminated. The remaining pixels are dilated to fill in any small gaps in the cell boundaries. Remaining pixels clusters that are not part of the cell boundaries are all located in small pixel clusters. It was found that eliminating clusters less than 50 pixels left all of the cell boundaries intact. The result is shown as the center picture in Figure 5. The last step is to eliminate all remaining clusters that do not coincide with the high standard deviation clusters of Figure 4. This last figure clearly shows that many of the cells are missing at least part of their cell boundary definition.

Many of the cells on these images are dead cells and can be identified by a round appearance, bright interior, and dark boundary. These cells are important to identify as a normal terminal fate in cell development and differentiation. Some of the round cells are cells that are not yet attached to other cells. However it is not important to track the random movement of the dead cells once they come in contact with other cells. Many of the dead cells have brighter than average

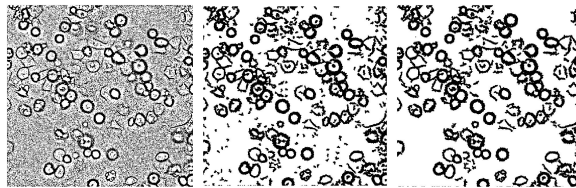


Fig. 5. Steps to finding the dark cell borders.

pixels scattered just inside the boundary. To enhance these bright regions, pixels are selected that are at least two standard deviations above the mean value for the whole image. These bright pixels are randomly located across the image, but are seen in very small clusters of 3 or more pixels near cell borders. To enhance these bright regions, we dilate only the bright pixels in these clusters, and not the bright pixels scattered elsewhere. We evaluate the shapes of the resulting clusters of bright pixels, and the round clusters, as well as those completely surrounded by the dark regions, are initially assigned cell numbers. The resulting cells for the above image are shown in Figure 6.

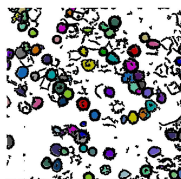


Fig. 6. Bright round cells are assigned first.

The next step is to identify the cells that do not have clear cell boundaries on the image. Potential cell pixels are selected by looking for a combination of pixels that are at least slightly above the mean pixel intensity, and also pixels whose standard deviation value is at least the mean standard deviation value. Several steps to assemble these pixels are shown for this example image in Figure 7. On the image these cells are recognized by the human eye because they are slightly brighter than the background. So we collect pixels slightly above the mean image intensity value (a value of 100 is used as a cutoff because it is approximately 1/3 of a standard deviation higher). Again to pick up the small clusters and leave out the background noise, we selectively dilate pixel clusters of greater than 3 pixels. Then pixels with standard deviation values less than the mean standard deviation are eliminated. Remaining pixels are dilated again, but cannot grow into the dark regions.

Of the remaining pixels, we look at regions that have not been previously assigned to a cell above, and keep clusters bigger than 100 pixels, a cutoff for the size of a cell. We use the pattern of the pixel standard deviations from Figure 4

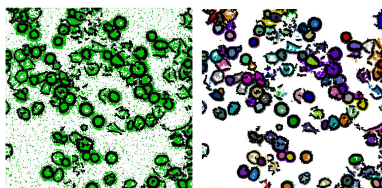


Fig. 7. Pixels with intensity above 100 are identified and clusters greater than 3 are dilated. Resulting pixels are kept if the standard deviation is at least the mean standard deviation of the image (om).

to understand which of the resulting pieces are part of the same cell. The pixels with the very highest standard deviation values are either part of the dark border of a cell and its surrounding pixels, or part of the very brightest cell pieces. If two clusters of this set fall along the same high standard deviation region, they are probably both part of the same cell. The very high standard deviation pixels, defined by pixels whose standard deviation is at least the mean (om) plus two overall standard deviations ($2*osd$) are shown in Figure 8, along with the larger than 100 cluster pieces and the final grouping of these pieces. A few final steps check the resulting cells, fill any small holes, and eliminate cells that are too small.

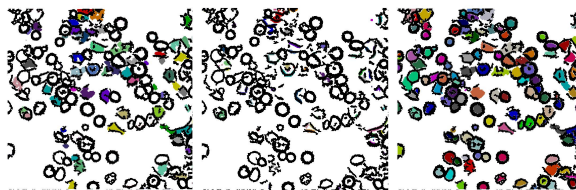


Fig. 8. Clusters larger than 100 pixels (left) are combined according to regions of very high standard deviation (center), and the final initial segmentation is shown (right).

6 Simultaneous Segmentation and Cell Tracking

Once an initial segmentation is performed, subsequent images can use information from the previous image to fill in gaps in cell boundaries, where they may not be present in an image. The beginning steps, defining a standard deviation value for each pixel, defining dark regions where they are present around the cells, and finding the brightest, round cells, are the same as those described above. Cell numbers are assigned to the bright round cells according to their location and size. A cell that moves significantly between images is matched by brightness and size across steps. As an example, Figure 9 shows a section of three

consecutive images, where one of the cells is seen to be moving rapidly towards the bottom of the image.

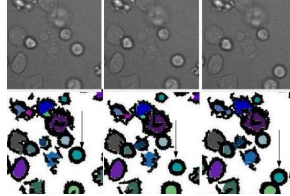


Fig. 9. Section of three consecutive images showing one of the cells (noted by the arrows) moving rapidly towards the bottom of the image and the corresponding sections of the segmentation masks.

Once the bright round cells are numbered, the rest of the cells are then defined by pixel intensity standard deviation, areas brighter than the mean, sections of dark boundaries, and boundaries markers from a previous image. The steps are shown below. Figure 10 first shows the initial cell numbering for the bright, round cells on the left. Pixels representing the bright high standard deviation areas, or pixels associated with a cell in a previous image are shown on the right. As described above, pixels with intensities greater than 100 are collected, and clusters of greater than three pixels are selectively dilated. The resulting pixels are then kept if they have higher than the mean value of standard deviation, or if they are at locations of pixels from cells of the previous image. At this stage the clusters are split by the available dark boundaries and numbered. Clusters of pixels that are associated with a single cell from the previous image are assigned to that cell number. Clusters of pixels associated with more than one previous cell are eroded so that they split into several sections, each associated with separate cells or at least fewer cells. The eroding process is repeated if necessary. Figure 11 shows the clusters that include more than one cell, and then a final eroded set of clusters.

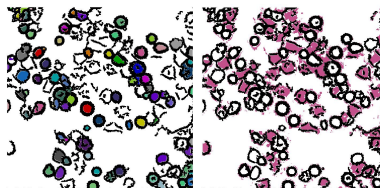


Fig. 10. Numbered round bright cells (left), pixels that are brighter than average, have a high standard deviation, or are in locations of pixels from a previous cell image (right).

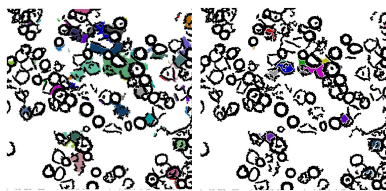


Fig. 11. Clusters of pixels associated with more than one cell. Several of these clusters are eroded to define separate cell clusters.

There are a few final steps where accuracy is checked. Bright shadows under dark regions that look like potential new cells are eliminated. Gaps in cell boundaries that are overestimated in areas with bright shadows are redefined using the cell boundaries of the previous image. Cells that are less than 30 pixels are eliminated. Figure 12 shows the segmentation as a function of time for the first 30 minutes of cell growth. The images and resulting masks are given at 10-minute intervals. This segmentation method can keep track of mitosis events, with an example shown in Figure 13, where a section of images taken two minutes apart shows a dividing cell, and the corresponding segmentation masks also show that event.

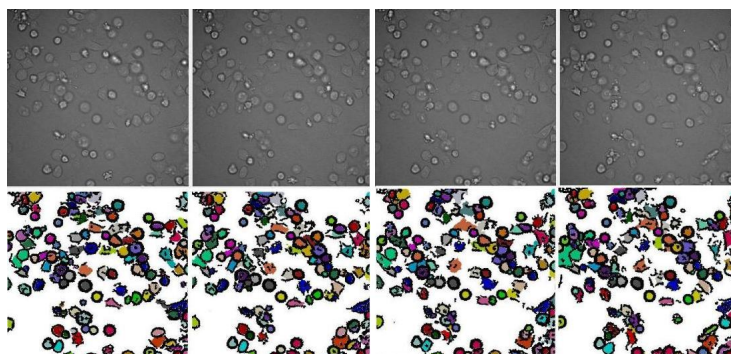


Fig. 12. Images and resulting masks at 0, 10, 20, and 30 minutes.

7 Conclusions

For initial growth of this breast cancer cell line in culture, our segmentation/cell tracking program is capable of handling the five required tasks to robustly analyze migration of cells over time: segmenting and identifying cells, individually naming cells, assigning cells of one image to corresponding cells in the preceding and succeeding images, identifying dead cells, and identifying mitosis events.

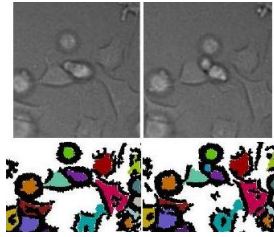


Fig. 13. Dividing cell at center of section of images taken two minutes apart, and the resulting segmentation masks.

For short periods we are able to track every cell in our test set of images. The purpose of this paper is to present new techniques for the challenging segmentation issues this cell line presents. Groundtruth data for this cell is painstakingly hard to collect for all of the cells, but future work will include larger scale tests with this method and comparisons with limited manually collected data. Future work will also include cell shape analysis to maintain high accuracy for longer periods. Because we track more than just a center cell location, we are able to analyze size and shape change information from these data over time as well, which will provide further insight into the biology of these cells.

8 Acknowledgements

This research was supported in part by the Intramural Research Program of the National Cancer Institute, NIH, Bethesda, USA.

References

1. Chaffer, C.L., Weinberg, R.A.: A perspective on cancer cell metastasis. *Science*; 331, 1559–64 (Mar 25, 2011)
2. Miller, F.R., Santner, S.J., Tait, L., Dawson, P.J.: MCF10DCIS.com xenograft model of human comedo ductal carcinoma in situ. *J Natl Cancer Inst*, 92, 1185–6 (2000)
3. Simon, I., Pound, C.R., Parin, A.W., Clemmes, J.Q., Christens-Barry, W.A.: Automated Image Analysis System for Detecting Boundaries of Live Prostate Cancer Cells. *Cytometry*, 31, 287–294 (1998)
4. Tscherepanow M, Zollner F, Hillebrand M, Kummert F: Automatic Segmentation of Unstained Living Cells in Bright-Field Microscope Images. In *International Conference on Mass-Data Analysis of Images and Signals in Medicine, Biotechnology, Chemistry and Food Industry (MDA)*. 5108, 158–172, Springer (2008)
5. Zhang, K., Xiong, H., Yang, L., Zhou, X.: A Novel Coarse-to-Fine Adaptation Segmentation Approach for Cellular Image Analysis. *Advances in Multimedia Modeling*, 4351–4362 (2007)
6. Zanella, C., Campana, M., Rizzi, B., Melani, C., Sanguinetti, G., Bourguin, P., Mikula, K., Peyrieras, N., Sarit, A.: Cells Segmentation from 3-D Confocal Images of early Zebrafish Embryogenesis. *IEEE Transactions on Image Processing* (September, 2009)

7. Wang, M., Zhou, X., Li, F., Huckins, J., King, R.W.: Novel cell segmentation and online SVM for cell cycle phase identification in automated microscopy. *Bioinformatics*, 24(1) 94101 (2007)
8. Palaniappan, K., Ersoy, I., Nath, S.: Moving object segmentation using the flux tensor for biological video microscopy. *LNCS -Proc. PCM 2007*, Springer, 483–493 (2007)
9. Stuelten, C.H., Busch, J.I., Tang, B., et al.: Transient tumor-fibroblast interactions increase tumor cell malignancy by a TGF-Beta mediated mechanism in a mouse xenograft model of breast cancer. *PLoS One*, 5, e9832 (2010)
10. Tang, B., Vu, M., Booker, T., et al.: TGF-beta switches from tumor suppressor to prometastatic factor in a model of breast cancer progression. *JClinInvest*, 112, 1116–24 (2003)
11. Friedman, N., Russell, S.: Image segmentation in video sequences: a probabilistic approach. *Proceedings 13th Conference Uncertainty Artificial Intelligence*, (1997)
12. Kachouie, N.N., Fieguth, P., Ramunas, J., Jervis, E.: A Statistical Thresholding Method for Cell Tracking. *IEEE International Symposium on Signal Processing and Information Technology*, (2006)