

Automated Quantification of Colony Growth in Clonogenic Assays

Kishore Mosaliganti¹, Jian Chen², Firdaus Janoos¹, Raghu Machiraju¹, Weiming Xia², Xiaoyin Xu³, Kun Huang⁴

¹Department of Computer Science and Engineering, The Ohio State University, Columbus, OH, USA

²Center for Neurologic Disease, ³Functional and Molecular Imaging Center, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

⁴Department of Biomedical Informatics, The Ohio State University, Columbus, OH, USA

Abstract—The need to determine the behavior of cell colonies in cultures often arises in controlled biological experiments. The deployment of microassays involving multiple wells allows the assessment of parameters such as the stage of malignancy, growth constraints and chemical concentrations etc. The large amount of preprocessed data negates the use of manual counting methods and requires automatic image analysis algorithms. In this work, we determine the clonal growth patterns in mutated cell lines as part of a drug development study. After suitable preprocessing, individual cells are segmented out in each frame independently. In a given frame, the cells are grouped into clonal colonies by estimating their geodesic distance to colony boundaries available in the previous frame. The geodesic distance metric is defined using the statistical 2-point correlation functions (2-pcfs). Therefore, we progressively track the clonal colony growth in an Eulerian framework by utilizing cell segmentations in conjunction with the 2-pcfs. We report our results on sample culture cells and validate the same with manual measurements.

Index Terms—Phase-contrast microscopy, clonal colonies, cell populations, image segmentation, N -point correlation functions

I. INTRODUCTION

It is quite well established that tumors have a clonal origin – all of the affected and transformed cells can be traced to a small set of mutated precursor cells that lose the ability to undergo apoptosis (programmed cell death) and hence continue to proliferate [1]. In in-vitro cell cultures, it is observed that mutated cells tend to form colonies due to rapid proliferation [2]. These colonies contain closely packed cells that expand to occupy the entire surface area. Clonogenic assays containing several wells are a popular form of measuring the impact of different types of toxic chemicals, mutagens, carcinogens and radiation on cells [3]. In such studies, a large number of experiments are carried out to understand the effect of experimental parameters such as frequency of chemical or radiation exposure and chemical concentrations etc. Therefore, manual counting methods have been rendered impractical owing to their subjectivity and cumbersome process [3]–[5]. There is a growing need for automated, objective and robust methods for quantification analysis that can perform cellular profiling (segmentation) and temporal tracking.

As a consequence of developing novel therapies for cancer, our collaborators and co-authors (Chen and Jia, Harvard Medical School) are interested in determining the effect of certain chemicals in preventing or reducing the colony growth in mutated cells. Hence, normal cells in culture are exposed to

radiation to undergo mutation and then treated with a chemical. The wells are then observed using phase-contrast microscopy at various fixed time-points. *The goal of image analysis is to track the growth of cell colonies (not individual cells) across the frames and detect colonies formed in the later frames.* The framework needs to be robust to noise arising from acquisition and other sources.

In phase-contrast microscopy, cells are detected based on variations of refractive index in the path of light. The images are usually of low contrast and contain background inhomogeneities. Cellular organelles often cause cell interiors to have a spotted appearance obfuscating image gradient field. We use an image processing pipeline that facilitates foreground extraction and estimation of image gradients. Also, we wish to point out that cells lying on colony boundaries are in contact with other cells (different colonies) in later frames. A cell is the basic unit of a clonal colony and hence, in order to separate merged colonies, cell segmentation is a necessary preprocessing prerequisite.

Figure 1 show a sequence of three frames in a typical dataset. While the individual clonal colony regions are easily marked out in the first frame, the task is increasingly complicated in the later frames. The cells need to be grouped into colonies that existed in the previous frame based on their *proximity* in location and prevalent cellular spatial arrangement. In this context, the 2-point correlation functions (2-pcfs) serve to estimate cellular packing densities and spatial distributions. The 2-pcfs forms a Riemannian manifold to evaluate the geodesic proximity of a cell to the colony boundary. In other words, a Voronoi partitioning on the 2-pcf manifold is achieved to classify the cells. Once all the cells associated with a given colony are identified, we employ the 2-pcfs once again in defining the new colony boundary repeating the process in the following frame. Thus, we track individual colonies. It should be noted that we essentially propose an Eulerian framework to monitor and track cell colonies.

We applied our methods on 144 data sets each containing three temporal frames. Manual validation of a dataset having 21 colonies is obtained by individually marking the clonal contours. Sensitivity and specificity ratios of our framework is computed using the manual validation as ground-truth.

The rest of the paper is organized as follows. In Section II, we describe related work in cell segmentation, N -pcfs and temporal tracking. Section III motivates and explains different stages of the image processing pipeline. We especially describe the role of 2-point correlation functions in providing a suitable

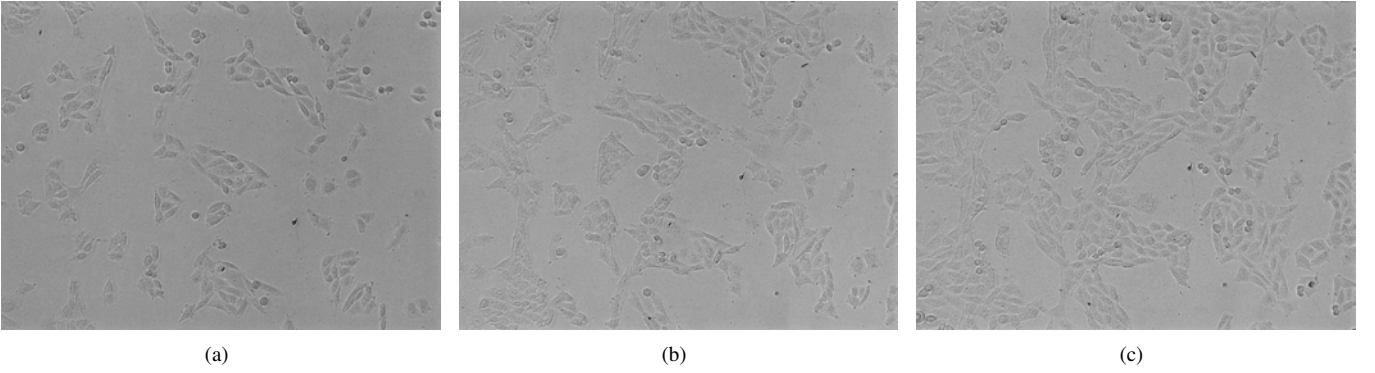


Fig. 1. A sequence of three frames captured at (a) 0 hour (b) 12 hour (c) 24 hours. The frames display a rapid proliferation of cells and merging of cell colonies.

metric function to evaluate geodesic distances. In Section IV, the tracking framework is formally developed while section V reports on the implementation and the results obtained. Finally, in Section VI, we provide a summary and describe our plans for the future.

II. RELATED WORK

A vast amount of research is conducted to analyze temporal cell behavior and attributes such as migration, proliferation, reproduction cycle, differentiation, apoptosis and quiescence. Example of such studies exist in cancer biology [2], cellular pathology [3], [4], [6], biomaterial design [7], [8], stem cell research [9] and genetic phenotyping [10]. The goals of image analysis include cell segmentation, colony segmentation and tracking in an automated and objective fashion.

A preliminary task in image analysis is to first extract the relevant cell/tissue foreground region. For example, in [5], cell foreground is identified in each frame using a Bayesian method that inspects the histograms of the foreground and background. Barber *et al.* [3] use image morphological operators to detect/refine the cell foreground. Global methods fail when deployed on our images owing to image inhomogeneities. In our case, we adaptively threshold preprocessed images to obtain the cell foreground [11]. In microscopy images, cells often appear as touching or overlapping with each other. There exist methods which deal with separating the cells [12]–[14]. In [15], we employ the geodesic active contours level-set method to split cells that elegantly incorporates the overlapping neck cues, image gradients and a cell shape model.

Ultimately, we seek to identify clonal colonies in our image frames for tracking purposes. We adapt methods from material science literature to obtain segmentations of colonies [16]. Microstructure irrespective of its origin (material science or biology) may be defined as a collection (ensemble) of points, lines, internal surfaces, and volumes [17]. Statistical distributions of such geometric attributes of ensemble features collectively specify the geometric state of a microstructure.

Existing methods for automated object tracking may be classified as belonging to two different approaches: tracking by model evolution and tracking by detection [18]. In the first approach, the object-of-interest is first detected in each frame and a correspondence problem is solved across the

frames [19]. This approach is most suitable when the objects are well-separated and faces problems if the objects maintain close-contact or may merge and split. The second approach involves the creation of mathematical shape models that are evolved over time to follow object movements. Parametric active contours [20], mean-shift models [21], [22] and level-set evolution [23], [24] fall under this category. Our approach does not track shapes. It is Eulerian in essence as opposed to most methods that are Lagrangian. We determine attributes of the domain that best describe cell colonies.

Li *et al.* [5] used a level-set approach to track migrating and proliferating cells that were imaged with phase-contrast microscopy. They imposed topological constraints on level-set evolution to mimic cell-like behavior (for example, cells never merge). In our case, we are interested in tracking an entire clonal colony and not individual cells. The three frames afforded to us are exactly sufficient to track colonies but not individual cells.

III. IMAGE ACQUISITION AND PROCESSING

In this section, we describe the dataset (Section III-A) and an image pre-processing pipeline. This is followed by a brief introduction to the N -point correlation functions that are used in defining a geodesic distance metric in the next section.

A. Dataset

The images were acquired by General Electric IN CELL Analyzer 1000 microscope with a 10X objective. The microscope has a motorized stage that allows it to automatically take brightfield and fluorescence images over time in a raster scanning pattern. The images were collected over cultured cells that were placed in 96-well plates. Images were collected every twelve hours at 0 hour, 12 hours and 24 hours. Each frame had image dimensions of approximately 1000×1000 (on average). The nature of the imaging process, long time gap between imaging periods and changing cell demographics introduce the following artifacts that require preprocessing steps:

- 1) Figure 2a shows an example of the shot noise arising from the presence of cellular organelle structures in the cell. These structures possess different refractive indices

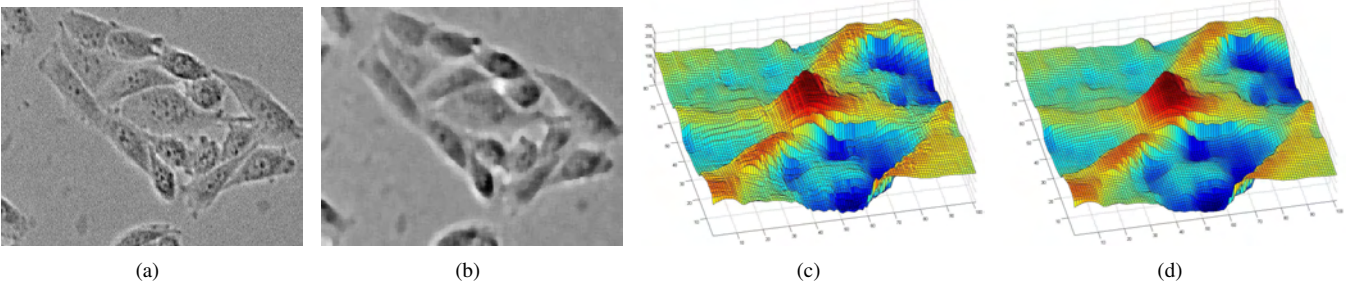


Fig. 2. A zoomed portion of frame 1 is shown. (a) The micrograph depicts a colony of cells with shot-noise arising from cellular organelles within. (b) The same image after using median filtering techniques. (c) An intensity map of a small region that shows the high frequency noise. (d) After anisotropic diffusion filtering, the intensity profiles are more smooth.

for light transmission and therefore show up as unwanted structures in the image. Since we are interested in cells and populations, these structures tend to obfuscate local gradients that in turn, complicate the cell segmentation problem. They are removed by applying a non-linear gray-scale median filter of sufficient width (7×7) that preserves larger structures. Figure 2b shows the image after eliminating shot noise.

- 2) The image intensity exhibits noise and inconsistent variations in contrast, especially along the cell boundaries. Figure 2c and show a small portion of an image, and the corresponding intensity map. The lack of smoothness in the intensity field restricts us from obtaining meaningful cell segmentations (the boundary edges provide important cues). We use the non-linear diffusion method [25] to smooth noise, while preserving cell boundaries (Figure 2d).
- 3) Image inhomogeneities (background gradients) complicate the setting of global parameters. For example, the threshold to extract image foreground varies across the frame. It also causes background fragments to be identified as foreground. We employ adaptive thresholding scheme to address this problem (Section III-B).
- 4) Cells form the basic unit of a colony. In our tracking analysis, we try to group cells into salient colonies. However, cells often appear as overlapping or touching each other in the foreground mask and hence need to be individually segmented. In Section III-C, we employ a geodesic level-set algorithm in segmenting cells and thereby address these problems.
- 5) Some mutated cells undergo apoptosis and are no longer relevant to the analysis. They shrink in size, have a compact shape and exhibit a weaker dark intensity profile. Using the cell segmentations, we process the images to detect dead cells and remove them from the cell populations.

B. Adaptive Thresholding

Earlier, we described the image inhomogeneities that exist in our images. Figure 3a shows an example where the background gradients are bad. A single threshold when applied on the entire image provides an inconsistent result. We use adaptive thresholds that are determined from local histogram

medians [11]. We first convolve the grayscale image I with the median filter of sufficient width (30×30). The window size is chosen to be statistically relevant to the size of a single cell. The original image is subtracted from the convolved image (I_W) as $D = I - I_W$. Finally, we threshold the difference image D with a user-defined threshold. If the pixel value is below the threshold it is set to the background value, otherwise it assumes the foreground value.

Using connected component analysis, we measure the size and shape of each component. Components smaller than a user-defined threshold in size and with anisotropic aspect ratios are also eliminated from consideration since they most likely belong to the background. Often, the cell foreground contains small islands of background internally. Such holes are also filled up to account as cellular foreground. We have now estimated the cellular foreground image F (Figure 3b).

C. Cell Segmentation

The cell foreground mask F does not present segmentations of each individual cell. We use the level set methods for front propagation as a numerical technique for obtaining an initial segmentation of the cells [15] within the mask. Cells are essentially separated on the basis of the well-defined gradient around their boundaries as well as by the presence of a short-neck region between two overlapping cells. The two key steps for the level set method are:

- 1) *Embedding the surface*: we represent the cell boundary $\Gamma(t)$ as the zero level-set of the signed distance function $\psi(\mathbf{x}, t)$. Formally, $\Gamma(t) = \{\mathbf{x} : \psi(\mathbf{x}, t) = 0\}$.
- 2) *Embedding the motion*: we need to derive the update equation 1 such that the motion of the zero level set has the desired properties described below. For this purpose, we use the active contour formulation developed by Caselles *et al.* [26]:

$$\frac{\partial \psi}{\partial t} = g(B)(c + \kappa)|\nabla \psi| + \nabla g(B) \cdot \nabla \psi \quad (1)$$

The weighting factor g refers to the reciprocal function of the cell boundary information B ($g(B) = \frac{1}{1+B^2}$). The boundary information B is comprised of the image gradients and *neck* information. The first speed term consists of the curvature and propagation dependent functions. The parameter κ is the curvature along the normal to the level-set contour and c is a balloon force that is added to evolve the curve

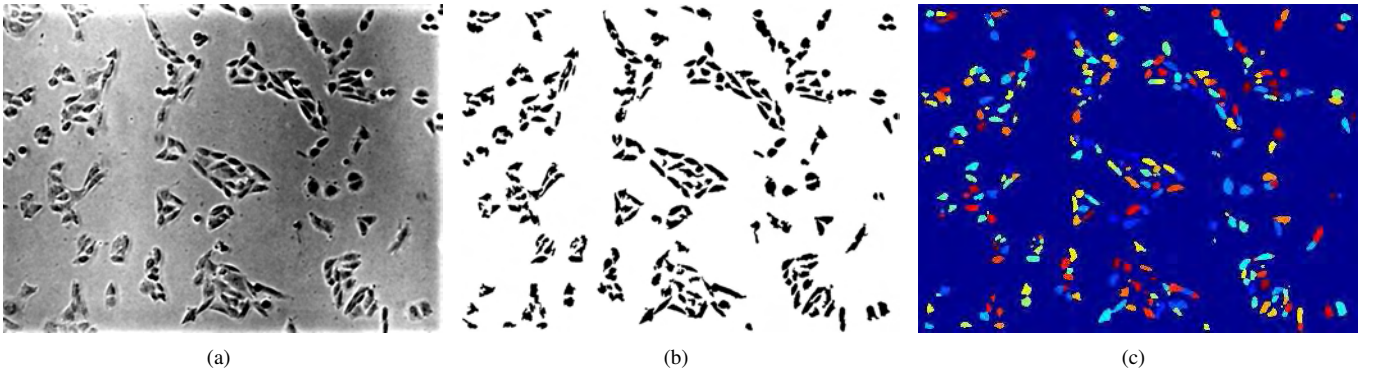


Fig. 3. (a) Figure shows frame 1 from Figure 1 after applying the preprocessing techniques and histogram equalization. It is easy to notice the image inhomogeneities that prevent the application of global thresholding strategies for foreground extraction. (b) After applying adaptive thresholding. (c) Cellular segmentation from the foreground map in (b) and the gradient map of (a).

outwards. Note that the boundary is characterized as a local minima in the metric function g . The term $\nabla g \cdot \nabla \psi$ constitutes the boundary attraction term. Hence the contour lies on the cell boundary upon convergence (Figure 3).

Speed functions $g(B)$: Let $D(F)$ and $D(\tilde{F})$ represent the unsigned distance fields emanating from the contour of the cellular foreground in F lying outside and within the cell foreground respectively. Consider the image $P = D(F) + D(\tilde{F})$. $D(\tilde{F})$ causes the appearance of a directional minima in the *neck region* represented by a blue line segment Figure 4a. The appearance of a directional minima is a direct consequence of the neck shape resulting in lower magnitudes of the distance field. Similarly, the cell-background interface is also marked by a well-defined local minima. In fact, the cell boundary is marked by a value of 0 in P by definition, and positive elsewhere. Both these reasons contribute in *halting* the evolution that results in cell segmentation. While the presentation so far accounts for the neck cues, we also observe that some cells exhibit high gradients at the site of overlap. As is standard practice, we take the image gradients (∇I) into account. Hence, our boundary information B is now defined as:

$$B = \frac{\nabla I + \alpha(D(F) + D(\tilde{F}))}{1 + \alpha} \quad (2)$$

α represents a scaling factor ranging from $[0, \infty)$. For large values, the neck cues are assigned more weight than the inter-cellular gradients and vice-versa. The level-set initialization occurs inside each cell. Hence we use the foreground image $D(\tilde{F})$ and choose maxima in the distance field as our initialization. Figure 4 show a pair of overlapping cells, the speed function $g(B)$ and the final segmentations obtained as a result.

D. 2-Point correlation functions

A clonal colony consists of closely packed cells with low separation distance in the image domain. Using the cell segmentations in each frame, we group cells into existing clonal colonies. This requires us to develop a notion of distance of a cell from a clonal colony. The geodesic distance between a cell and the colony should be low if the cell belongs to the colony and vice-versa. Earlier, we mentioned about the similarity in cellular arrangements with material microstructure. Based on

this observation, we utilize the N -point correlation functions (N -pcfs) to define a distance metric. These functions measure the cellular packing densities and evaluate to a high value in the interior of a colony region and diminish at the boundaries. The 2-pcfs ($N = 2$) have been initially introduced in medical imaging for histology image segmentation [27], [28].

Let $F : \Omega \rightarrow \{0, 1\}$ be a binary image of the cellular (1) and background (0) phases. A 2-pcf $P_{i_1 i_2, \ell}$ is the probability that a straight line segment of length ℓ when randomly placed in the microstructure has one vertex in phase $i_1 \in \{0, 1\}$ and the other vertex in phase $i_2 \in \{0, 1\}$. For a 2-phase microstructure, there are four possible 2-pcfs namely $P_{00, \ell}$, $P_{01, \ell}$, $P_{10, \ell}$ and $P_{11, \ell}$ and:

$$P_{00, \ell} + P_{01, \ell} + P_{10, \ell} + P_{11, \ell} = 1 \quad (3)$$

To estimate this function at a pixel location $p \in \Omega$, we resort to using Monte Carlo methods to sample the distributions of cells in a local neighborhood ω (typically 50×50) of p . We place a number of randomly oriented and positioned line segments typically of length $\ell = 30$ in ω . We then count the fraction of line segments that have an end point in phase i and the other in j to give an estimate of the 2-pcf $P_{ij, \ell}(p)$.

The 2-pcf estimates four independent probabilities $\{P_{00, \ell}, P_{01, \ell}, P_{10, \ell}, P_{11, \ell}\}$ that each capture different information. Now, $P_{00, \ell}$ measures the probability of the background phase being present at both vertices. Hence, interior of the background regions are marked with values closer to 1. Similarly, $P_{11, \ell}$ evaluates to values closer to 1 in the interior of the cellular foreground. The functions $P_{01, \ell}$ and $P_{10, \ell}$ possess higher values on the boundaries of the cellular and background phases. Hence, regions with a probability values around 0.5 capture the boundaries. For tracking colonies, we require a probability map that differentiates colony from non-colony regions. It is defined as:

$$\mathcal{P} = (1 - P_{00, \ell}) \times |P_{00, \ell} - 0.5| \times |P_{00, \ell} - 0.5| \times P_{11, \ell} \quad (4)$$

Figure 6(a-c) shows the probability maps \mathcal{P}^1 , \mathcal{P}^2 and \mathcal{P}^3 computed on the frames from Figure 1.

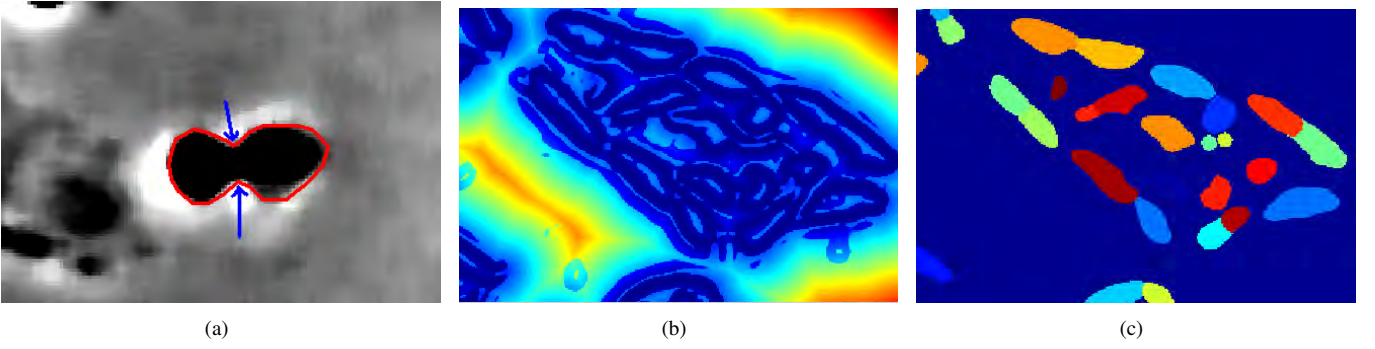


Fig. 4. (a) A pair of overlapping cells. The *neck* region is indicated by blue arrows. (b) Speed map $g(B)$ of the zoomed portion of cells from Figure 2. It is easy to notice the role of gradients and the distance field in creating minima at cell overlap regions. (c) Cell segmentations obtained after splitting overlapping cells.

IV. COLONY TRACKING

Let $F^i : \Omega \rightarrow \{0, 1\}$, $I^i : \Omega \rightarrow \mathbb{N}$ and $\mathcal{P}^i : \Omega \rightarrow [0, 1]$ represent the i^{th} frame containing cell foreground, cell segmentations and the 2-pcf evaluation for $i \in \{0, 1, 2\}$. Further, let \mathcal{C}_k^i be the centroid of the k^{th} cell in I^i . We represent δ_j^i to be the boundary of the clonal colony $j \in \{1, 2, \dots, N\}$ in frame i .

A. Distance metric

We now introduce a Riemannian metric \mathbf{G} defined in terms of the 2-pcf probability \mathcal{P} evaluated on the image F^i as:

$$\mathbf{G} = \nabla \mathcal{P}(F^i) \nabla \mathcal{P}(F^i)^T \quad (5)$$

Infinitesimal distances under \mathbf{G} are measured by:

$$\|dx\|_{\mathbf{G}}^2 = dx^T \mathbf{G} dx = (dx^T \nabla \mathcal{P}(F^i))^2 \quad (6)$$

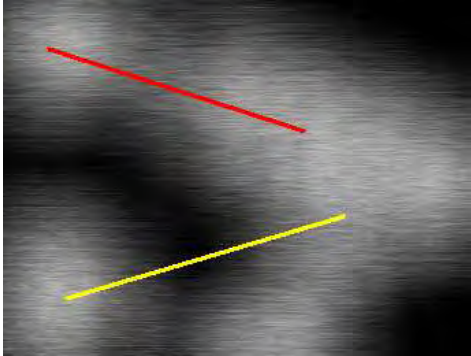


Fig. 5. Using our $N\text{-pcf}$ \mathcal{P} based distance metric, the distance along the red curve and yellow curve will be quite different although the curves are of the same length in the domain Ω . Since the red curve lies within the same clonal colony, the distance is relatively low. The yellow curve forms a bridge between two different colonies and has to pass through the background phase in between.

Under this metric, distances measured parallel to the large gradients are higher and vice-versa. Figure 5 provides a simple illustration. For example, consider any pair of cell centroid \mathcal{C}_j^i and \mathcal{C}_k^i separated by a Euclidean distance d in the image domain Ω and lying within the same clonal colony. Then, the 2-pcf will evaluate to high values in the shortest path from $u \rightarrow v$ and present a low gradient along the path. Hence,

the distance between them will be small. However, if \mathcal{C}_j^i and \mathcal{C}_k^i belong to different clonal colonies, the intermediate region presents high gradients in the 2-pcf values and hence evaluates to a much larger distance.

The solution to the correspondence problem determines the affiliation of each cell k in frame i into one of the N clonal colonies. The classification is achieved by determining the colony boundary ($\delta_{j^*}^i$) nearest to the given cell (abstracted by its centroid \mathcal{C}_k^i).

$$\langle j_k^{i*} \rangle = \arg \min_{j \in \{1, 2, \dots, N\}} d(\mathcal{C}_k^i, \delta_j^{i-1}) \quad (7)$$

B. Colony Segmentation using Fast Marching Methods

In Equation 7, the distance of a given cell \mathcal{C}_k^i to all the colony boundaries $\Delta = \{\delta_j^{i-1} | j < N\}$ is calculated and the colony at a minimum distance is chosen as its affiliation. This is equivalent to finding generalized Voronoi partitions of the 2-pcf manifold given Δ as its generating set. Therefore, we use the fast marching methods in evolving the boundary contours with the 2-pcf (\mathcal{P}^i) as its contour speed. We note the arrival time of each contour at individual cells and the earliest arriving contour is set as the colony affiliation of that cell.

The fast marching methods consists of an PDE formulation that is implemented using narrow-band numerical schemes thereby making it efficient. Let $T(p) : \Omega \rightarrow \mathbb{N}$ be the arrival time of the front as it crosses $p \in \Omega$.

$$|\nabla T_j^i| \mathcal{P}^i = 1 \quad T_j^i = 0 \text{ on } \delta_j^i \quad (8)$$

The calculation of the colony affiliation for a cell in Equation 7 can now be posed in our new formulation as:

$$\langle j_k^{i*} \rangle = \arg \min_{j \in \{1, 2, \dots, N\}} T_j^i(\mathcal{C}_k^i) \quad (9)$$

C. Re-initializing Clonal Boundaries

In a given frame i , after computing all the cells that constitute a clonal colony j , we need to determine the new colony boundary δ_j^i . We proceed as follows:

- (i) Compute the 2-pcf function on an image containing cells belonging to colony j alone. The function evaluates to a high value in regions where the cell density is high and vice-versa.
- (ii) Select a suitable threshold $T = 0.5$ on the function values

that provides us with a binary mask of the colony region.

(iii) A simple canny edge detection filter gives us δ_j^i . This new boundary is then repeated to perform the tracking in frame $i + 1$.

V. IMPLEMENTATION AND RESULTS

This section reports on a number of experiments we conducted to evaluate our framework for colony tracking. Experiments were conducted with a validated dataset having 3 frames. Our goals were three-fold: (i) We demonstrate the ability of the framework to robustly track colonies in the presence of acquisition noise. (ii) The proposed framework is computationally efficient for processing large microscopy images. (iii) We also show that real datasets benefit from the availability of automated tools for objective analysis.

In order to validate our algorithm, 3 images were reviewed and all pertinent colonies were roughly delineated with boundaries. These boundaries are not exactly drawn around a clonal population but only serve to encapsulate all the relevant cells alone within the boundary. This defines our ground-truth to compare our algorithm results. We admit that more data is needed. We earlier mentioned that cells form the basic unit of a colony. Hence, we only need to compare the algorithm detected colony cells and the ground-truth colony cells.

The matched pairs can be separated and counted in 4 categories: (i) a counts the number of *true-positive* results. (ii) b counts the number of *false-positive* results. (iii) c counts the number of *false-negative* results. (iv) d counts the number of *true-negative* results. Thus, one can define the sensitivity and specificity of our framework as follows:

$$\text{sensitivity} = \frac{a}{a + c} \quad \text{specificity} = \frac{d}{b + d} \quad (10)$$

$$\text{error} = \frac{b + c}{a + b + c + d} \quad (11)$$

Sensitivity measures the effectiveness of the algorithm in detecting the true-positives. A high value indicates that it is easy for the algorithm to rule out a spurious match with a high probability. On the other hand, a value of high specificity, can identify a *bona fide* match with a high probability of success.

Frame	Cell Counts	Sens.	Spec.	Error
Frame 1	378	0.9134	0.8827	0.0215
Frame 2	610	0.8407	0.8222	0.123
Frame 3	794	0.7807	0.8022	0.1733

TABLE I

SENSITIVITY AND SPECIFICITY VALUES FOR COLONY TRACKING

Sensitivity and specificity values for the colonies detected in the 3 frames is given in Table I. We observe that the framework has a good sensitivity (> 0.84), specificity (> 0.83) and with a low error rate (< 0.15) on average. It is easy to see that the performance of the algorithm is best in the earlier frames and relatively diminishes in later frames. This is expected since the extensive clonal growth creates merged boundaries. In turn, this leads to the cells lying on the Voronoi boundaries to be easily affected. We also observed in several instances that a

single colony in the starting frame emerged as two separate colonies in later frames. As a result of which the tracking step combined the two colonies into one of the same affiliation. As part of our future work, we plan to extend our algorithm to account for new colonies emerging in later frames.

All our experiments were compiled and executed on an Intel P4 3.2 GHz processor equipped with 1GB RAM. The computation of the cell segmentation algorithm [15], *2-pcf* algorithm [27] and fast marching level-set methods were developed in C++ using the *The Insight Segmentation and Registration Toolkit* (ITK) libraries [29]. The preprocessing framework was developed using the MATLAB Image Processing Toolbox. In Table II, we provide the stage-wise execution time in seconds for the three frames. Our algorithms required under 10 minutes per frame. Each frame had dimensions of approximately 1000×1000 pixels on average.

Frame	Cell-Segmentation	2-pcf	Colony-Tracking
Frame 1	198	221	411
Frame 2	241	215	423
Frame 3	311	274	476

TABLE II

EXECUTION TIME IN SECONDS FOR THE 3 FRAMES

VI. CONCLUSIONS

In this paper, we described methods to temporally track the evolution of clonal colonies across frames. A segmentation of the individual cells is achieved in each frame by using standard algorithms from literature. The cells are grouped into clonal colonies by noting their geodesic distance from the clonal colony boundary established in the previous frame. The geodesic distance metric is suitably defined using the 2-point correlation functions (*2-pcfs*). The *2-pcfs* are novel neighborhood estimators of cellular packing and density distributions. The framework was successful in matching all corresponding clonal colonies. Using available ground-truth, the sensitivity of our matching algorithm was measured at $> 84\%$ with a specificity of $> 83\%$. In future, we shall perform extensive validation and incorporate the detection on new colonies in later frames. We also propose to explore the higher-order correlation functions towards segmentation and tracking applications. The *2-pcf* estimators can complement other tracking algorithms and work in a real-time scenario.

REFERENCES

- [1] P. Fialkow, "Clonal origin of human tumors," *Annual Review of Medicine*, vol. 30, pp. 135–143, 1979.
- [2] A. Yao and H. Rubin, "Automatic enumeration and characterization of heterogeneous clonal progression in cell transformation," *Proceedings of the National Academy of Sciences*, vol. 90, pp. 10524–10528, 1993.
- [3] P. Barber, B. Vojnovic, J. Kelly, C. Mayes, P. Boulton, M. Woodcock, and M. Joiner, "Automated counting of mammalian cells," *Physics in Medicine and Biology*, vol. 46, pp. 63–76, 2001.
- [4] J. Dahle, M. Kakar, H. Steen, and O. Kaalhus, "Automated counting of mammalian cells by means of a flat bed scanner and image processing," *Cytometry*, vol. 60A, pp. 63–76, 2004.
- [5] K. Li, E. Miller, L. Weiss, P. Campbell, and T. Kanade, "Online tracking of migrating and proliferating cells imaged with phase-contrast microscopy," in *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW '06)*, June 2006, pp. 65 – 72.

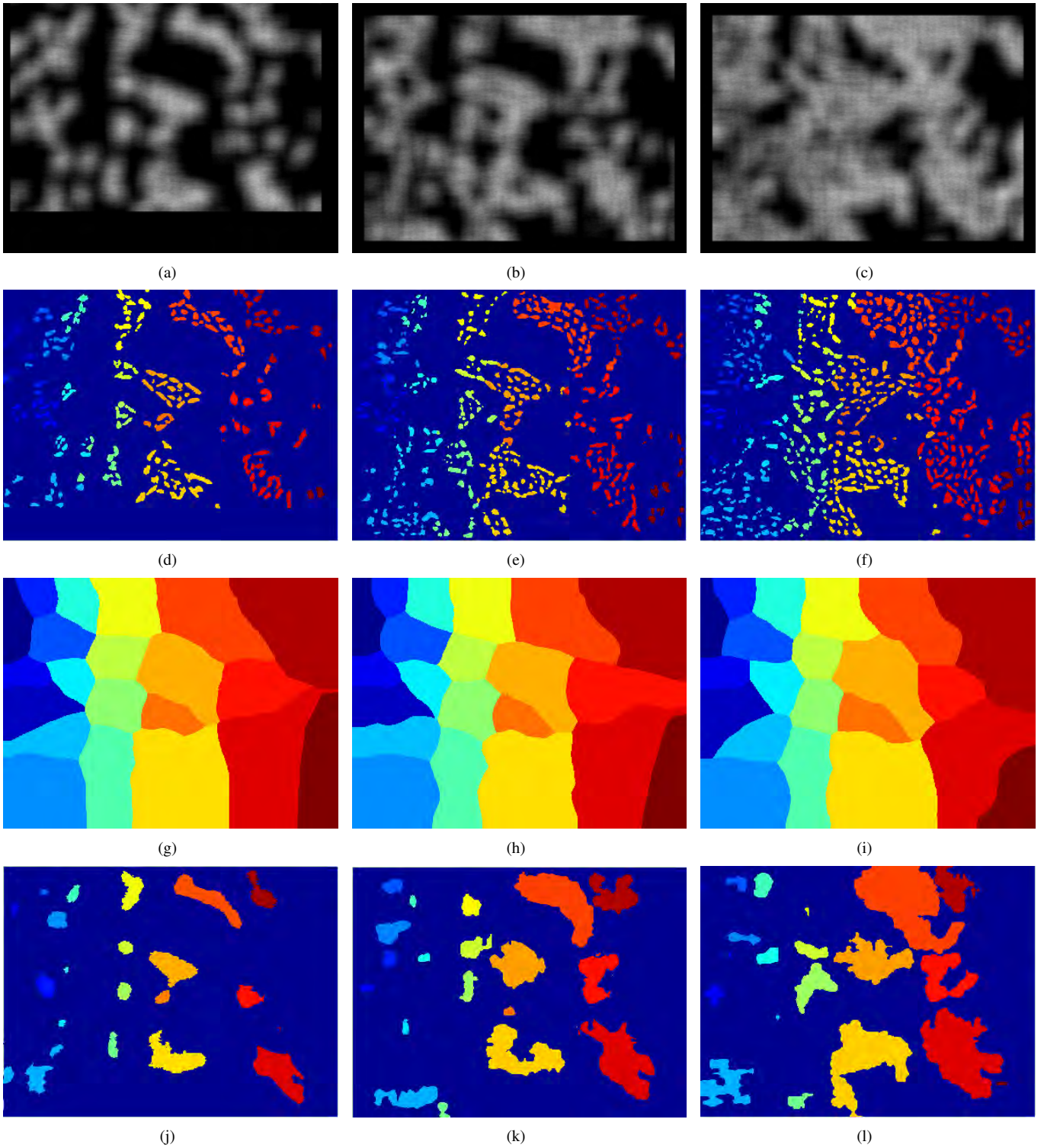


Fig. 6. The 2-pcf \mathcal{P} when computed at (a) 0 hour (b) 12 hour (c) 24 hour is shown for the frames in Figure 1. Cellular segmentations of the frames using the methods in Section III-C is shown for the same frames in (d-e-f). The Voronoi partitions of image domain Ω using the fast marching methods described in Section IV-B for the same frames (g-h-i). The fast marching method for a given frame uses the colony segmentations (j-k-l) from the previous frame in conjunction with the corresponding 2-pcf in (a-b-c). New segmentations of the colonies are obtained by using the 2-pcf from Section IV-C.

- [6] P. Ranefall, L. Egevad, B. Nordin, and E. Bengtsson, "A new method for segmentation of color images applied to immunohistochemically stained cell nuclei," *Analytical Cellular Pathology*, vol. 15, no. 3, pp. 145–156, 1997.
- [7] E. Miller, G. Fisher, L. Weiss, L. Walker, and P. Campbell, "Dose-dependent cell growth in response to concentration modulated patterns of fgf-2 printed on fibrin," *Biomaterials*, vol. 27, no. 10, pp. 2213–2221, 2006.
- [8] P. Campbella, E. Miller, G. Fisher, L. Walker, and L. Weiss, "Engineered spatial patterns of FGF-2 immobilized on fibrin direct cell organization," *Biomaterials*, vol. 26, no. 33, pp. 6762–6770, 2005.
- [9] P. Wenzel, L. Wu, A. deBruin, W. Chen, G. Dureska, E. Sites, T. Pan, A. Sharma, K. Huang, R. Ridgway, K. Mosaliganti, R. Sharp, R. Machiraju, J. Saltz, H. Yamamoto, J. Cross, M. Robinson, and G. Leone, "Rb is critical in a mammalian tissue stem cell population," *Genes and Development*, vol. 21, no. 1, pp. 85–97, 2007.
- [10] L. Wu, A. de Bruin, H. I. Saavedra, M. Starovic, A. Trimboli, Y. Yang, J. Opavska, P. Wilson, J. Thompson, M. Ostrowski, T. Rosol, L. Woollett, M. Weinstein, J. Cross, M. Robinson, and G. Leone, "Extra-embryonic function of Rb is essential for embryonic development and viability," *Journal of Nature*, vol. 421, pp. 942–947, 2003.
- [11] F. Yan, H. Zhang, and C. Kube, "A multistage adaptive thresholding method," *Pattern Recogn. Lett.*, vol. 26, no. 8, pp. 1183–1191, 2005.
- [12] S. Beucher, "The watershed transformation applied to image segmentation," in *Conference on Signal and Image Processing in Microscopy and Microanalysis*, 1991, pp. 299–314.
- [13] T. Jones, A. Carpenter, and P. Golland, "Voronoi-based segmentation of cells on image manifolds," in *Proceedings of the ICCV Workshop on Computer Vision for Biomedical Image Applications (CVBIA)*, 2005, pp. 535–543.
- [14] M. Tscherepanow, F. Zöllner, and F. Kummert, "Automatic segmentation of unstained living cells in bright-field microscope images," in *Workshop on Mass-Data Analysis of Images and Signals in Medicine, Biotechnology and Chemistry MDA*, 2006, pp. 86–95.
- [15] K. Mosaliganti, L. Cooper, R. Machiraju, K. Huang, and G. Leone, "Visualization of cellular biology structures: When a little means a lot," in *Submitted to the IEEE Conference on Visualization*, 2007.
- [16] S. Torquato, *Random Heterogenous Material*, Springer Verlag, 2004.
- [17] A. Gokhale, A. Tewari, and H. Garmestani, "Constraints on microstructural two-point correlation functions," *Scripta Materialia*, vol. 53, pp. 989–993, 2005.
- [18] E. Meijering, I. Smal, and G. Danuser, "Tracking in molecular bioimaging," *Signal Processing Magazine, IEEE*, vol. 23, no. 3, pp. 46–53, 2006.
- [19] I. Smal, W. Niessen, and E. Meijering, "Particle filtering for multiple object tracking in molecular cell biology," in *IEEE Nonlinear Statistical Signal Processing Workshop*, 2006, pp. 44.1–44.4.
- [20] C. Zimmer and J. Olivio-Marin, "Coupled parametric active contours," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 11, pp. 1838–1842, 2005.
- [21] O. Debeir, P. Van Ham, R. Kiss, and C. Decaestecker, "Tracking of migrating cells under phase-contrast video microscopy with combined mean-shift processes," *IEEE Trans. Med. Imaging*, vol. 24, no. 6, pp. 697–711, 2005.
- [22] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 142–151.
- [23] O. Dzyubachyk, W. Niessen, and E. Meijering, "A variational model for level-set based cell tracking in time-lapse fluorescence microscopy images," in *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2007, pp. 97–100.
- [24] D. Mukherjee, "Level-set analysis of leucocyte detection and tracking," *IEEE Transactions on Image Processing*, vol. 13, pp. 562–572, 2006.
- [25] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM Journal of Numerical Analysis*, vol. 29, no. 1, pp. 182–193, 1992.
- [26] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal on Computer Vision*, vol. 22, no. 1, pp. 61–97, 1997.
- [27] R. Ridgway, O. Irfanoglu, R. Machiraju, and K. Huang, "Image segmentation with tensor-based classification of N-point correlation functions," in *MICCAI Workshop on Medical Image Analysis with Applications in Biology*, 2006.
- [28] F. Janoos, O. Irfanoglu, K. Mosaliganti, R. Machiraju, K. Huang, P. Wenzel, A. deBruin, and G. Leone, "Multi-resolution image segmentation using the 2-point correlation functions," in *IEEE International Symposium on Biomedical Imaging*, 2007.
- [29] L. Ibáñez and W. Schroeder, *The ITK Software Guide*, Kitware, Inc., 2003.