

Segmenting time-lapse phase contrast images of adjacent NIH 3T3 cells

J. CHALFOUN*, M. KOCIOLEK†, A. DIMA*, M. HALTER*,
A. CARDONE*, A. PESKIN‡, P. BAJCSY* & M. BRADY*

*National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.

†Technical University of Lodz, Lodz, Poland

‡National Institute of Standards and Technology, Boulder, Colorado, U.S.A.

Key words. Background reconstruction, cell population, live cell segmentation, mitotic cell detection, quantitative imaging, separating cells in contact.

Summary

We present a new method for segmenting phase contrast images of NIH 3T3 fibroblast cells that is accurate even when cells are physically in contact with each other. The problem of segmentation, when cells are in contact, poses a challenge to the accurate automation of cell counting, tracking and lineage modelling in cell biology. The segmentation method presented in this paper consists of (1) background reconstruction to obtain noise-free foreground pixels and (2) incorporation of biological insight about dividing and nondividing cells into the segmentation process to achieve reliable separation of foreground pixels defined as pixels associated with individual cells. The segmentation results for a time-lapse image stack were compared against 238 manually segmented images (8219 cells) provided by experts, which we consider as reference data. We chose two metrics to measure the accuracy of segmentation: the ‘Adjusted Rand Index’ which compares similarities at a pixel level between masks resulting from manual and automated segmentation, and the ‘Number of Cells per Field’ (NCF) which compares the number of cells identified in the field by manual versus automated analysis. Our results show that the automated segmentation compared to manual segmentation has an average adjusted rand index of 0.96 (1 being a perfect match), with a standard deviation of 0.03, and an average difference of the two numbers of cells per field equal to 5.39% with a standard deviation of 4.6%.

Introduction

Automated microscopy and image analysis of live cells is critical to studying cellular dynamics at the single cell level.

Correspondence to: Joe Chalfoun, National Institute of Standards and Technology, Gaithersburg, MD 20899, U.S.A. Tel: +301-975-3354; fax: +301-975-6097; e-mail: Joe.chalfoun@nist.gov

With time dependent, single cell data, it is possible to uncover temporal patterns that cannot be resolved by population-averaged measurements (Sigal *et al.*, 2006; Halter *et al.*, 2011). The quantification of cellular reporters such as green fluorescent protein (GFP) allows analysis of the kinetics and fluctuations associated with gene expression, promoter activation and protein synthesis (Langenbach *et al.*, 2006; Sigal *et al.*, 2006; Sigal *et al.*, 2006b; Halter *et al.*, 2009), but requires accurate segmentation of cells to enable accurate measures of cellular fluorescence.

Automated segmentation and tracking pose significant challenges due to the variability of image and object qualities. The variability arises from differences in imaging devices and modalities. In addition, different cell lines can have different physical characteristics and these characteristics can change in response to external conditions during a cell cycle. A particular challenge is accurately segmenting individual cells when cells are in contact with each other, particularly as the culture grows and cell density increases. The main motivation of this work is to address this problem and enable the collection of single cell data from cell monolayers as they approach confluence.

For this study, we have focused on segmenting live NIH 3T3 cells imaged by Zernike phase contrast microscopy. By segmenting phase contrast images, the fluorescence channels are available for quantifying expression from fluorescence protein reporters and the segmentation is determined in a manner that is not biased by intracellular location or intensity of fluorescent marker. Here, we present a new segmentation technique that consists of (1) background reconstruction to obtain foreground pixels with high signal-to-noise ratio and (2) incorporation of biological insight about dividing and nondividing cells into the segmentation algorithm. The latter aspect of the segmentation provides the basis for a more robust separation of foreground pixels of connected cells

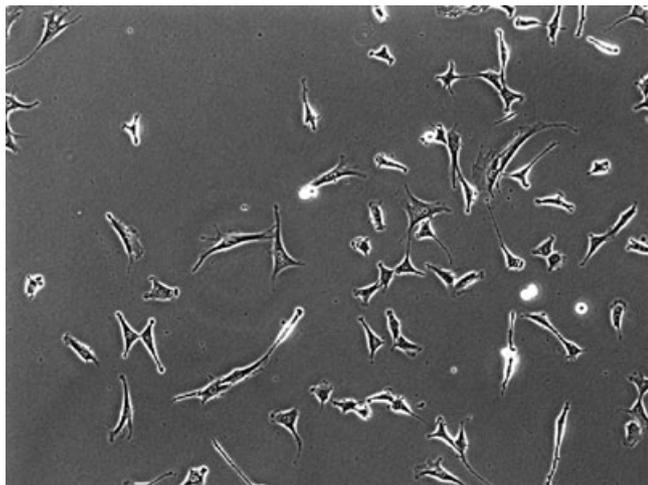


Fig. 1. Representative phase contrast image of NIH 3T3 cells approximately 24 h after seeding. The number of cells in contact increases throughout the time-lapse image set used in this study. The image shown is 696×520 pixels (approximately $1.4 \text{ mm} \times 1.0 \text{ mm}$).

into individual cell pixels, compared to methods that do not incorporate any prior biological knowledge.

Figure 1 shows an example of a NIH 3T3 live cell image where cells are in contact with each other. The visual cues in the image that are used to segment cells include (1) a bright halo region surrounding a cell in general, (2) a dark region inside these halos and (3) a circular bright region often found inside the dark central region where the nucleus of a cell is located. We segment the halo around the cell so that intracellular fluorescence can be accurately quantified (Halter *et al.*, 2011). By including the entire halo, the segmented cell area is enlarged which ensures that the entire cell is within the segmented mask. Intracellular fluorescence is measured by integrating the signal over the entire cell mask and subtracting a local background signal measured from around the cell periphery (Halter *et al.*, 2011). By including the halo, we reduce the probability that part of a cell is excluded from the foreground and included in the local background. When a cell is about to divide (cell undergoing a mitosis), it rounds up and becomes very bright under phase contrast. We have used these visual cues and biological understanding of these cells to automatically and accurately generate seeds for a watershed-based segmentation to separate adjacent cells.

We validate our segmentation technique by comparing our results against a reference segmentation input provided by experts for a stack of 238 images. The choice of the adjusted rand index (ARI) similarity metric is based on our desire to compare two mask images with similar overlapping areas but where the grouping of pixels into discrete objects may be different. Unlike overlap based similarity metrics, the ARI metric penalizes the result when groupings of pixels do not

correspond to the same superstructure. For example, if the reference mask image has one object in an area and the algorithm mask image has two or more objects in the same area, the ARI metric will be reduced because the number of object groups is different for the two images. A structured approach to the choice of the ARI metric for our study is described by Bajcsy *et al.* (Bajcsy *et al.*, 2012), where similarity metrics are recommended based on their sensitivity response to the user defined input of interest. We also chose another metric to compute the similarities at a cellular level: the number of cells per frame (NCF). The choice of this particular metric is based on its wide use as an indicator of cellular proliferation and cytotoxicity within a culture system.

Live cell image acquisition protocol

Before imaging, cells were seeded into a six-well plate at 1200 cells/cm^2 and allowed to adhere in the incubator for approximately 3 h. The cells were maintained throughout the imaging experiment in a humidified 5% CO_2 balanced-air atmosphere at 37°C using a microscope incubation chamber. Phase contrast and fluorescence images were acquired using an automated microscope controlled by 'ISee' image acquisition software ('ISee' Imaging Systems, Rayleigh, NC, USA; <http://www.iseeimaging.com/>). Images were acquired every 15 min for a time length greater than 62 h in 12 different fields using a $10\times$ (0.3 NA) objective on Zeiss 200M inverted microscope equipped with an automated stage (Ludl, Hawthorne, NY, USA), a collimated blue LED (470 nm) fluorescence excitation source (Thorlabs, Newton, NJ, USA) and a CoolSNAP HQ charge coupled device (CCD) camera (Roper Scientific, Tucson, AZ, USA). A $0.63\times$ de-magnifying lens was positioned in front of the CCD. Auto-focusing was performed during the acquisition and was limited to a $<5 \mu\text{m}$ z-axis sweep so that large focus adjustments that might lead to out-of-focus images were prohibited.

The raw image data from the detector are 696×520 pixels, and with a single channel values represented by unsigned numbers of 16 bits per pixel. The pixel to μm calibration was determined by using a graduated micrometre image. The pixel length was found to be $= 1.95 \mu\text{m}$ in each *x* and *y* direction. Further details describing the live cell imaging can be found in (Halter *et al.*, 2011).

Background reconstruction

The objective of background reconstruction is: (1) to support more accurate detection of foreground pixels (i.e., cell pixels) by subtracting the background signal and (2) to aid in detection of biological features used to separate cells in contact (see Section 4.2). The background reconstruction problem is a well-studied problem in surveillance, monitoring and tracking applications (Tang *et al.*, 2008; Kang *et al.*, 2009), but here, it is uniquely

applied to phase contrast images of cells, and contributes to the accuracy of the overall cell edge segmentation.

The following four criteria need to be met to apply our background reconstruction method: (1) an initial foreground/background mask can be calculated, (2) the CCD camera noise can be modelled by a stationary ergodic stochastic process (i.e., the statistical properties of the CCD camera noise do not change over time), (3) most cells do not stay in one location of the field of view (FOV) during the experiment and (4) for most pixels in FOV there is at least one time instance during the experiment when a pixel is not occupied by a cell. We show below that all of the 4 criteria were met with our test data set.

Initial foreground/background separation

The first step of the algorithm is the calculation of the absolute gradient image applying Robert's edge operator (Gonzalez & Woods, 2008; see Fig. 2a). This method performed very well in detecting most cells in these images.

The absolute gradient image is thresholded by using an automatically determined threshold value. The automated threshold selection assumes that there is sufficient number of background pixels in the image to produce high peak, similar to Gaussian distribution, in the part of the absolute gradient histogram with low intensity values. This assumption is satisfied in our example due to the low cell density at the beginning of the experiment. The threshold value is chosen as the bin value corresponding to the right tail around the peak of the gradient intensity distribution where the frequency value is approximately equal to the frequency value of the lowest gradient value in the image. The method for finding the automatic threshold value is shown in Figure 2(b).

After thresholding the image, the holes are filled by morphological operation and connected regions are labelled to obtain foreground and background masks. To remove all noncellular objects (like small debris) from the labelled image, we eliminate all regions smaller than the minimum cell area = 100 pixels. This value has been determined empirically in our study of NIH 3T3 cells. The resulting mask image is depicted in Figure 2(c). By inspection of this mask image it is clear that many cell areas were not detected by the initial segmentation technique. We apply a morphological dilation (Gonzalez & Woods, 2008) to the foreground/background mask to ensure that we exclude as many undetected cell pixels as possible without compromising background pixels. The size of a morphological dilation kernel depends on the accuracy of the initial foreground/background detection, and in our study the size was set to 10 pixels based on empirical determination. This value was sufficient to cover most cell pixels that were not initially detected as shown in Figure 2(d). The aforementioned software algorithms were implemented in Borland C/C++ and the executable can be downloaded for free from (Kocielek, 2009).

Determination of pixel values for the reconstructed background image

To determine the reconstructed background image, we calculate the mode (the value that occurs most frequently) for each location (x, y) from the FOV as described in Figures 3(a) and (b). The background intensity over time at a single, representative location (x, y) is plotted in Figure 3(b). Because the background intensities and the fluctuations in intensities at the same location are relatively stable through time, we conclude that the CCD camera noise can be modelled by a stationary ergodic stochastic process, thus demonstrating that our test data set meets the second criterion. In Figure 3(a) and in the plot in Figure 3(b), low intensity background pixel values are highlighted. We attribute these outliers to thin cellular extensions that appear dark (low intensity) in phase contrast. The initial segmentation incorrectly assigns pixels associated with the thin cellular extensions as background pixels. Because these outliers can reduce the time-averaged value (the mean) of a background pixel (x, y) , we use the mode instead, as a more robust pixel value in the reconstructed background image.

The image shown in Figure 3(c) indicates the frequency during which a pixel is considered as background pixel. From this image we concluded that most cells do not stay in one location of the FOV and at each location in the FOV the background was uncovered at least 15 times for this particular FOV, thus demonstrating that criteria 3 and 4 described above are met. However, even if these two assumptions were violated in some isolated pixel locations, the background value will be estimated by using the average of neighbouring background-reconstructed intensities. A representative reconstructed background image is shown in Figure 4.

In any experiment, it is not uncommon to get some images slightly out of focus. The impact of these blurred images is minimized by using the mode to compute the value of the reconstructed background pixels.

Cell segmentation

The previous work in the area of segmentation of phase contrast cell images can be classified into six method categories: (1) region growing methods [e.g., watershed algorithms (Roerdink & Meijster, 2001; Tscherepanow *et al.*, 2008)], (2) edge-based methods [e.g., Canny edge-based detection (Canny, 1986)], (3) energy function minimization methods (Williams & Shah, 1992; Dzyubachyk *et al.*, 2009), (4) cell motion based methods (Zanella *et al.*, 2010), (5) model-based methods (Kachouie *et al.*, 2007) and (6) methods based on advanced feature detection (Bradhurst *et al.*, 2008; Wang *et al.*, 2008)). Our segmentation method falls under the last category; it operates on a single image frame and uses the temporal information of live cell imaging only for background

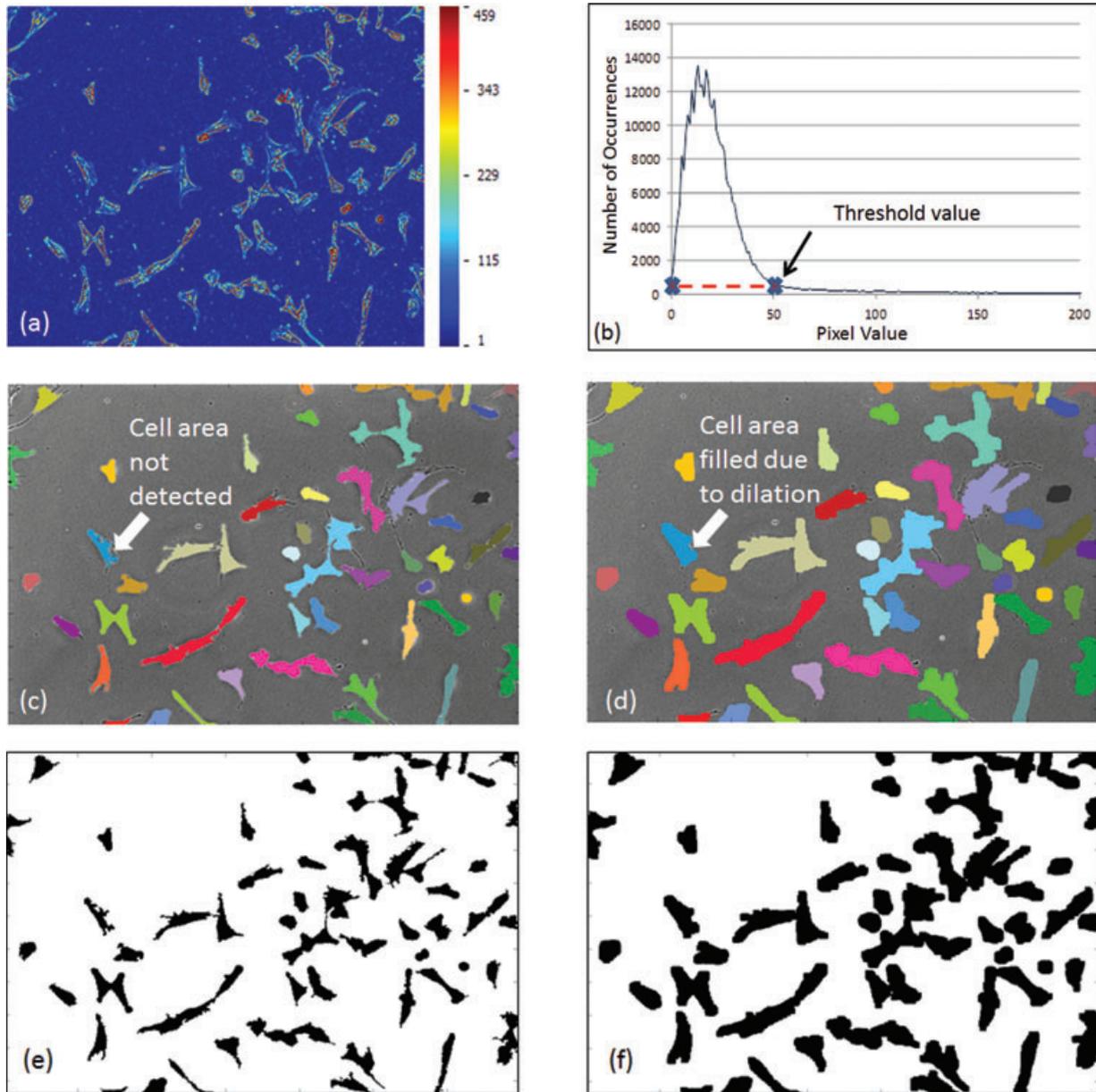


Fig. 2. Initial segmentation for identifying background image pixel values. (a) Absolute value of the gradient image shown in pseudo colours, corresponds to the raw image shown in Figure 1. (b) The histogram of the absolute values of the gradient image (a) for pixels values from 0 to 200. (c) Labeled image with holes in foreground removed. (d) The final foreground/background mask after morphological dilation. The white arrows in (c) and (d) indicate cell areas not detected by the initial segmentation but better included in the mask after morphological dilation. (e) The binary image that results after applying the threshold in (b) to the gradient image in (a). (f) The binary image after morphological dilation.

reconstruction. In this category, Wang *et al.* describe a series of feature-specific shape detection procedures to track cell cycle changes (Wang *et al.*, 2008). Ersoy *et al.* propose a flux-tensor based method, which detects cell movements (Ersoy & Palaniappan, 2008). Bise *et al.* (Bise *et al.*, 2009) uses temporal contour shapes of cells and tracking information to separate cluster of cells. In comparison to the related work, our study is unique in that we have explicitly employed biological insight to inform a segmentation process based on intensity and shape

characteristics. We show that incorporating the biological insight as rules in the algorithm improves the performance as compared with a relatively naive algorithm. In a similar vein, Bradhurst *et al.* (Bradhurst *et al.*, 2008) detect cell edges of bone marrow stromal cells in phase contrast microscopy images where the images are processed by rough and then refined segmentation. The work by Bradhurst *et al.* (Bradhurst *et al.*, 2008) uses biological knowledge of the images to overcome the bright halo surrounding stem cells and also pick up the

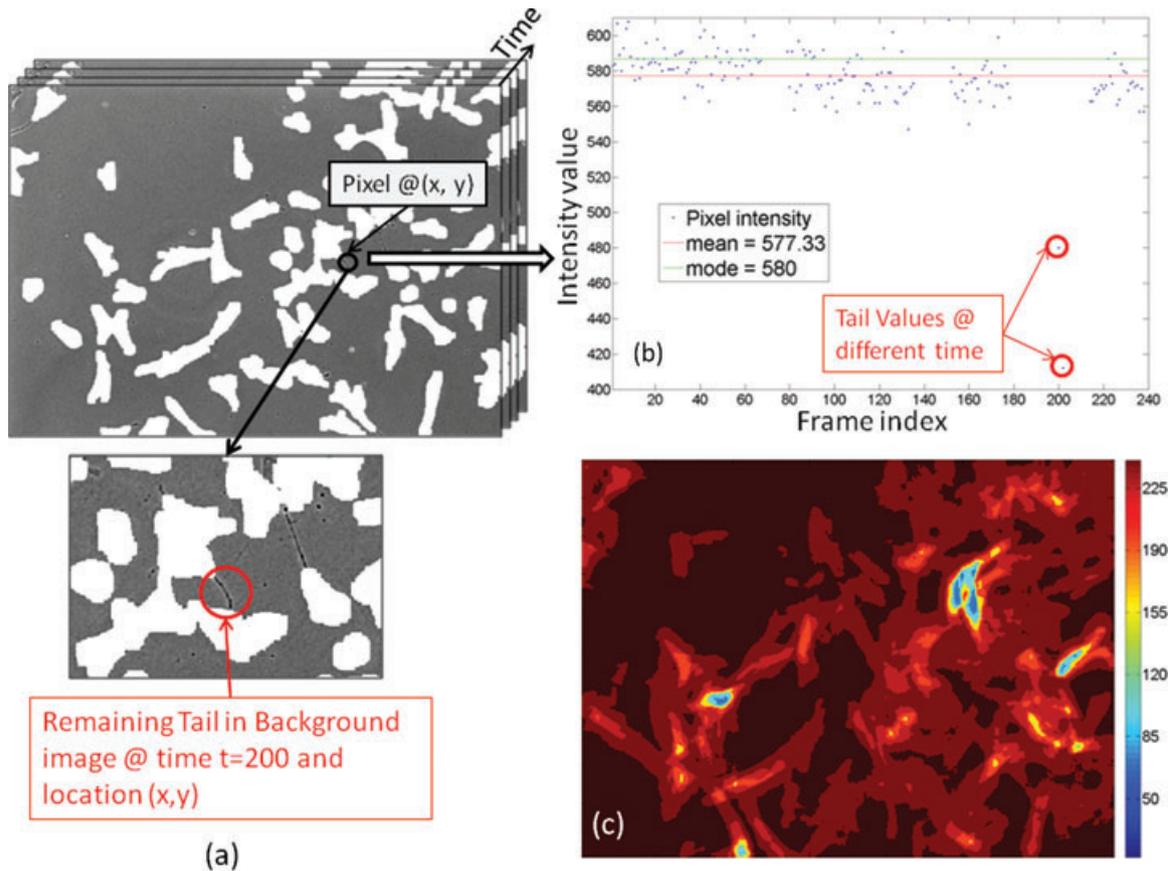


Fig. 3. (a) Image of cellular extension, indicated with the red circle in the enlarged image below, with low intensities remaining in background image. (b) Background pixel intensity plot at a single location (x, y) over time. The red arrows indicate low intensity values that can be attributed to the cell extension (tail) highlighted in the image in (a). (c) The frequency of occurrences of background pixels in the 238 time-lapse image set. Dark red indicates regions in the field of view rarely occupied by cells where a large number of observations are available to accurately calculate the background intensity.

dark interior region of cells. However, they do not address the issue of separating cells in contact.

The segmentation of adjacent cells is comprised of three steps: (1) foreground detection by background image

subtraction (Section 4.2), (2) foreground detection with biologically inspired algorithms (Section 4.2) and (3) separation of cells in contact (Section 4.3). All these steps are used to automate morphological seed selection. Morphological

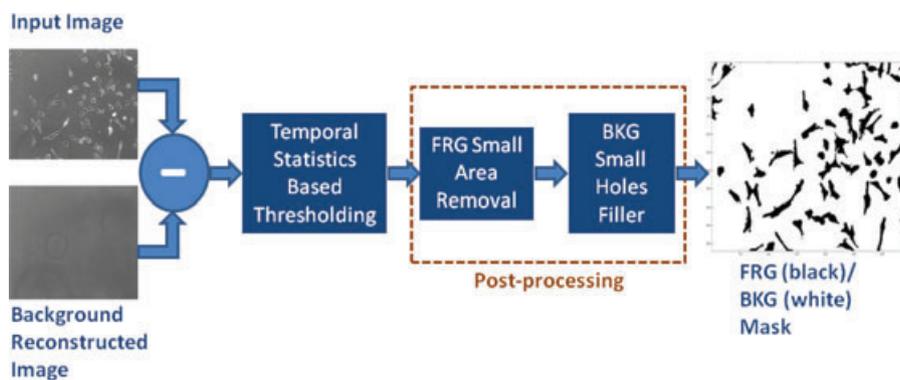


Fig. 4. Foreground detection steps. The 'Input Image' is the raw image from the CCD detector and the 'Background Reconstructed Image' is the computed from the time-lapse image set. Thresholding of the difference between the 'Input Image' and the 'Background Reconstructed Image' is followed by morphological operations to result in the foreground (FRG)/background (BKG) mask.

seeds are neighbourhoods or individual pixels from which a region-growing algorithm begins (Gonzalez & Woods, 2008). These seeds are used to minimize over-segmentation that typically comes from direct application of watershed transform (Debeir *et al.*, 2008).

Foreground detection by background image subtraction

To detect the foreground, we subtract the reconstructed background image from the phase contrast image. This yields image *I*1 with negative and positive values where negative values correspond to pixels darker than the background mode at that location and positive values correspond to pixels brighter than the background mode value. To obtain foreground pixels, we threshold the values at $\pm 2\sigma$ where σ is the sample standard deviation obtained at each pixel location and illustrated in Figure 3(b). This is a double thresholding of image *I*1 with the standard deviation matrix (SDM) calculated from the stack of time-lapse images. Each pixel (*x*, *y*) of the (SDM) matrix is equal to the standard deviation of all the background pixel intensities at that location. In this way, the threshold is determined statistically from the time-lapse image data. This step also assumes Gaussian distribution of background intensities that would assign background (BKG) and foreground (FRG) pixels correctly with a probability of 95.45%. Figure 4 shows a schematic of these processing steps including additional postprocessing operations that attempt to correct the misclassified pixels by considering a foreground region's minimum area and by filling background holes.

To quantify the performance of segmentation using background reconstruction, we computed the Dice index (Zou *et al.*, 2006) between masks coming from the expert segmentation and the ones coming from automated segmentation. The Dice index measures spatial overlap between two segmentations using the following formula $\text{Dice} = 2 \times \text{overlap} / (\text{area}_1 + \text{area}_2)$. We compare those results with the initial segmentation results described in Section 3.1 (automated threshold of gradient image using Robert's operator). The choice of the Dice index, over the ARI metric for this case, came from the fact that we are comparing binary images and not labelled ones with the interest being cell area segmentation. Thus, overlap-based similarity metrics are best suited for this type of comparison.

Figure 5 shows the difference between the initial segmentation and the background reconstruction. This plot highlights the improvement made by segmenting the cell edges (foreground/background binary mask) using the reconstructed background image. On average over 238 images the initial segmentation had a good agreement with the reference masks with a Dice index of 0.76 and a standard deviation of 0.036. However, when using the reconstructed background mask, Dice index shows a much better agreement between automated and reference masks with an average of

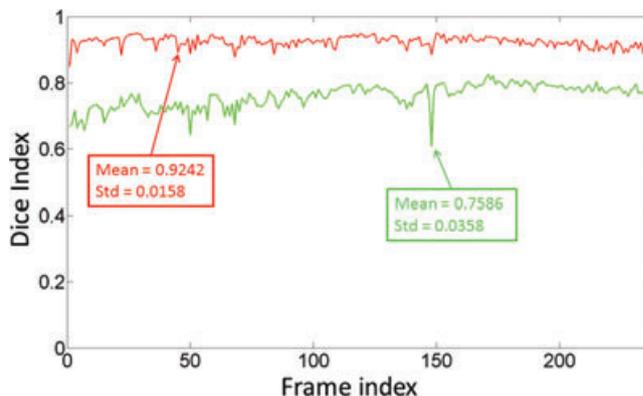


Fig. 5. Dice, an area similarity metric, computed between (1) reference data and segmentation with background reconstruction (in red) and (2) reference data and segmentation using initial segmentation without background reconstruction (automated threshold of gradient image using Robert's operator, Figure 2; in green).

0.92 and a standard deviation of 0.016. This implies that on average using the reconstructed background image for segmenting foreground pixels is 16% more accurate than using a threshold on the gradient of the image. This figure also proves that the initial segmentation and the segmentation using the reconstructed background are both very accurate on detecting the cell area in the phase contrast images due to the strong agreement between the binary masks coming from automated segmentation and the reference ones [a Dice index > 0.7 indicates excellent agreement between segmentation masks and reference data (Zijdenbos *et al.*, 1994)].

Foreground detection with biologically inspired algorithms

We have used our understanding of cell biology and visual cues in Zernike phase contrast images to develop a computational model for separating cells in contact with each other. The computational model, summarized in Table 1, consists of mapping biological states to image intensity and cell shape characteristics present in the images. When time-lapse, Zernike phase contrast images, of NIH 3T3 fibroblast seeded at low density are observed, a cell that divides appears to round up, split into two cells and re-spread on the surface. In the phase contrast images, the rounded dividing cell becomes very bright, likely due to the combination of the change in thickness and shape of the cell. Hence, the pixels corresponding to a dividing cell will be brighter than the pixels that belong to the nuclear region of the nondividing cell. We also observe that, in phase contrast images, the centre of nondividing 3T3 cells (nucleus location) is often brighter than the pixels over the cytoplasm of the cell. We use this information to derive a strategy for separating adjacent cells (Fig. 6). Each of the next two Sub-sections 4.2.1 and 4.2.2 describes an algorithm and the output of each algorithm is a mask image. These two mask images are used in Section 4.3 to separate cells in contact.

Table 1. Mapping biological and phase contrast imaging phenomena to computational models.

Cell state \ models	Biological model	Visual cues in phase contrast microscopy images	Computational model
Dividing cells	Cells release cellular attachments, depolymerize the cytoskeletal components and round up during cell division	1. Dividing cells are close to round in comparison with nondividing cell shapes	1. Roundness is measured by $R = 4\pi \times \text{area}/(\text{perimeter})^2$. Remove shapes where $R < 0.6$
		2. Dividing cells are associated with very high intensity pixels	2. Image intensity is thresholded at a value higher than the mean intensity of positive pixels over cell area
Nondividing cells	Cells are attached and spread	Interior of nondividing cells contains pixels with smaller or equal intensity values than the background (~dark cytoplasm and slightly brighter nucleus)	Image intensity is thresholded at a value lower than the mean intensity of positive pixels over cell area

Model for detecting dividing cells. To accurately distinguish the dividing cells in the image, the pixels with high positive intensities need to be detected and a roundness check applied to the resulting mask. Hence, we threshold the image after the background has been subtracted at an intensity equal to the mean of the resulting positive pixels. However, some of these high positive pixels will correspond to the halo around cells and these particular pixels need to be eliminated from the image. Because dividing cells are bright and round, eroding the resulting image, by five pixels in our study, will delete the pixels that correspond to the halo around cells and keep the pixels that correspond to dividing cells (see Fig. 6, upper right image). Finally, the nonround objects will not be selected. The roundness is measured by $R = 4\pi \times \text{area}/(\text{perimeter})^2$ and any object in the image will be removed if its roundness is less than 0.6 (user defined parameter).

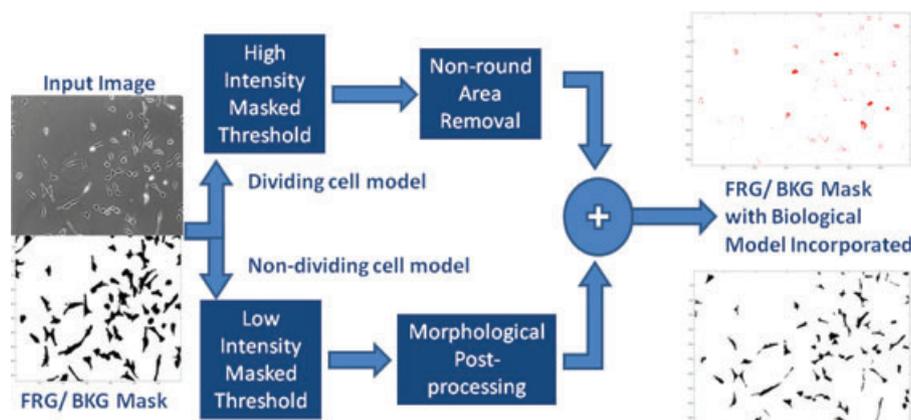
Model for detecting nondividing cells. To better separate adjacent cells, the dark cell interior region will be used as indicator of the approximate number of cells in a cluster. The dark cell regions are found using the original image with the

background subtracted. We again threshold at the mean of the positive pixels, but this time take all of the pixels below this value. Then erode the image and fill the holes. The pixels in the mask will, in general, correspond to a region containing cell nuclei (see Fig. 6, lower right image). Therefore, this serves as an indicator of the number of cells in a cluster.

These two biologically driven masks (dividing cell and the nondividing cell mask) will be used as seeds to aid the watershed algorithm to separate clusters of cells with minimized over-segmentation.

Separation of cells in contact

The Euclidian distance transform (Maurer & Qi, 2003; explained below) is applied to the original binary foreground mask as shown in Figure 7. The biologically driven masks are used to modify the distance transform by taking the distance values from the original distance transform and adding to it the value of the corresponding pixels that belong to the obtained masks.

**Fig. 6.** The workflow for executing the computational model for detecting foreground pixels corresponding to dividing cells and nondividing cells.

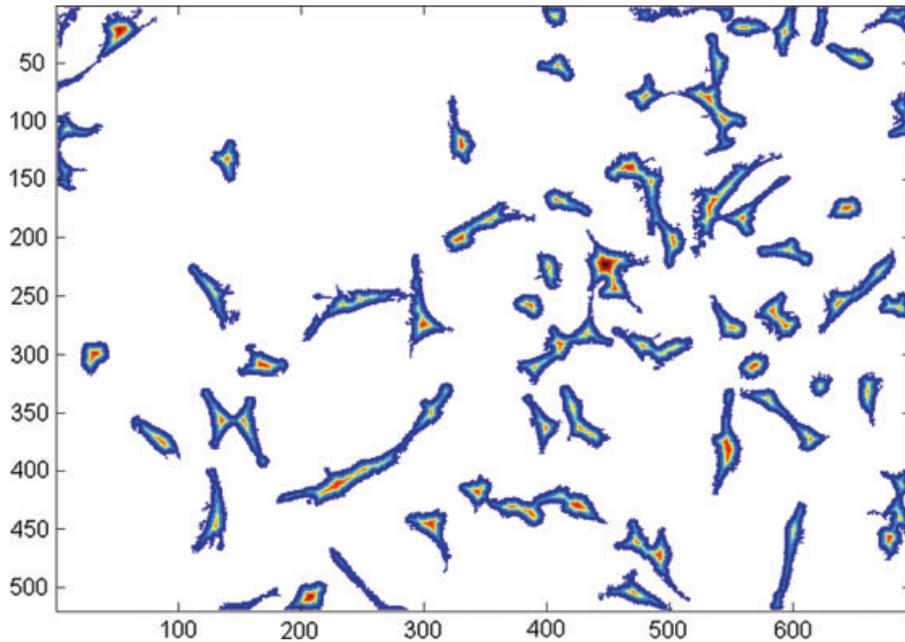


Fig. 7. The mask resulting from the initial segmentation (Figure 2) transformed by the Euclidean distance transform. The pixel intensity increases linearly with distance from the edge of the mask. The colour is proportional to the distance value with blue representing the minimum (value of 1) and the largest values are represented with red. In combination with mask images obtained by thresholding bright pixels and dark pixels in the raw image data, this transformed image is used to separate cells in contact.

The Euclidean distance transform of a binary 2D image I labels each foreground pixel (the foreground value is 1) with the distance of this pixel to the nearest background pixel (the background value is 0). For all foreground pixels of I , the algorithm determines the minimum Euclidean distance to the background pixels of I .

Let F be the set of n foreground pixels, $F = \{p \in I | I(p) = 1\}$, and B the set of m background pixels, $B = \{q \in I | I(q) = 0\}$, such that $F \cup B = I$ and $F \cap B = \emptyset$.

The new image DT is obtained from the original image I as follows:

$$DT(p) = \min_k \{d(p, q_k) | p \in F, q_k \in B, k = 1, \dots, m\} \quad (1)$$

$$d(p, q_k) = \sqrt{(x_p - x_{q_k})^2 + (y_p - y_{q_k})^2} \quad (2)$$

$$DT(p) = 0, \forall p \in B \quad (3)$$

Let DV be the mask of dividing cells and NV the one of nondividing cells. The following is true for every pixel: $DV \cap NV = \emptyset$.

We construct a modified distance transform image, MDT as follows:

$$\forall p \in DT | p \neq 0, MDT(p) = DT(p) + DT(p)^*(NV(p) \vee DV(p)), \quad (4)$$

where \vee is the logical OR operator.

This image captures the shape of a foreground region (the first term) and the biological knowledge about dividing cells

(round shape and bright cell intensity) and nondividing cells (dark intensity of a cell interior) in the NV and DV masks.

To segment and separate adjacent cells, we apply a modified watershed algorithm described in (Smolka, 2005) to the modified distance map, MDT. The resulting segmented image can be seen in Figure 8 where most adjacent cells are separated. One can visually see the difference between the images obtained with/without incorporating the biological features (Fig. 8).

Reference data-based comparison

The current bio-motivated algorithm is compared against the background subtraction algorithm to quantify performance improvement due to the prior biological knowledge embedded into the algorithm. When operating on cells that are in contact, general segmentation algorithms, such as Otsu, IsoData, Maximum Entropy, or our background subtraction algorithm, either over-segment cells into too many fragments or fail to separate cells. We have analysed the general segmentation algorithms in the past in Dima *et al.* (Dima *et al.*, 2011). We are not including the comparison of the bio-motivated algorithm against all the general segmentation algorithms in (Dima *et al.*, 2011) because they are conceptually similar and produce similar results as our background subtraction algorithm. Our validation approach is based on comparing the segmentation results against manually derived reference data. The reference data is derived from a time-lapse sequence of phase contrast

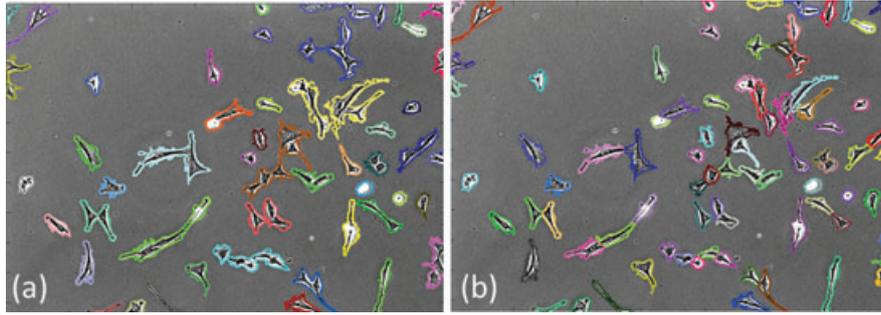


Fig. 8. (a) Foreground/background mask segmentation after background reconstruction before applying algorithms to separate adjacent cells in contact and (b) the final segmented image indicating improved cell separation of adjacent cells in contact.

microscopy images acquired at 15-min intervals for 60 h (238 frames total). The validation methodology consists of preparing reference data, defining segmentation accuracy metrics and computing accuracy between reference data and segmentation data.

Reference data preparation

A human expert manually segmented cells in each frame of the time-lapse images and separated adjacent cells by visual inspection: the expert took advantage of the time-lapse information to make a group of cell separation decision. The boundary is drawn using a computer mouse. This work was done using ImageJ software [<http://rsb.info.nih.gov/ij/>] and it took over 200 man-hours to complete the task. This reference data was inspected by a second expert to minimize human mistakes.

Segmentation accuracy metrics

We chose two metrics to evaluate the differences between the reference data and the automated segmentation results:

(1) ARI (Hubert & Arabie, 1985; Vinh *et al.* 2009) measures similarities between two segmented images (image1 and image2) at the pixel level. Let C_1 denote the group of labelled cells in image1 and C_2 the group of cells in image2. The number of cells in image1 and image2 is k_1 and k_2 , respectively. The ARI metric is based upon counting the pairs of pixels on which two cell labels in both images agree or disagree. ARI is bounded between 0 and 1 and is computed by the following formula:

$$ARI = \frac{\sum_{ij} \binom{n_{ij}}{2} - [\sum_i \binom{a_i}{2} \sum_j \binom{b_j}{2}]/\binom{T}{2}}{\frac{1}{2} [\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2}] - [\sum_i \binom{a_i}{2} \sum_j \binom{b_j}{2}]/\binom{T}{2}}, \quad (5)$$

$$\text{where } \binom{a}{b} = \frac{a!(b-a)!}{b!},$$

where T is the total number of data points, n_{ij} is the number of overlapping pixels between cell C_1 in image1 and cell C_2 in image2, $\binom{n_{ij}}{2}$ is a combination pair of data points, a_i and b_j are computed as follows:

$$a_i = \sum_{j=1}^{k_2} n_{ij} \text{ and } b_j = \sum_{i=1}^{k_1} n_{ij}$$

(2) NCF metric computes the differences between the number of correctly segmented cells per frame computed by detecting the under-segmented cells and the over-segmented cells and the total cell number in the masks generated by the expert in the reference data.

Evaluated segmentation methods

We evaluate two segmentation methods developed in this work to compare against the reference data set: (S1) segmentation based on the foreground mask without adjacent cell separation and (S2) biologically motivated segmentation with adjacent cell separation. In the reference data set, 8219 individual cell objects were identified throughout the image sequence containing 238 images. In comparison, the (S1) and (S2) segmentation methods identified 6382 and 7876 individual cell objects, respectively, throughout the time-lapse image set.

Experimental results

Figure 9 plots the ARI computed for the segmentation methods described above without and with adjacent cell separation (S1, red dashed line and S2, blue solid line) against the reference data. Figure 10 displays the comparison of S1 and S2 with respect to the reference data in terms of the NCF.

Discussion of experimental results

Figures 9 and 10 show that there is a high level of agreement between the reference set and both segmentations when cells occur in the FOV at very low density and the frequency of cell-to-cell contact is low. These low-density and low-interaction conditions are observed for image frames 1–150

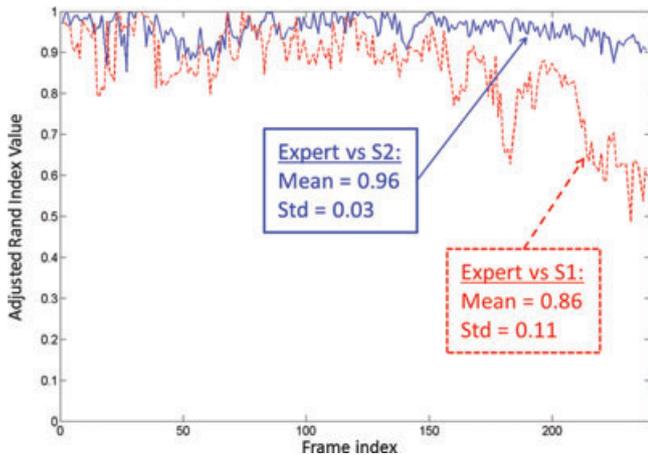


Fig. 9. Segmentation of single cells is improved and adjacent cells in contact are separated. Similarity measured using the ARI metric is plotted versus the frame index for the time-lapse image set for (1) Reference data and segmentation after applying algorithms to separate adjacent cells (S2) and (S2) Reference data and segmentation before applying algorithms to separate adjacent cells (S1).

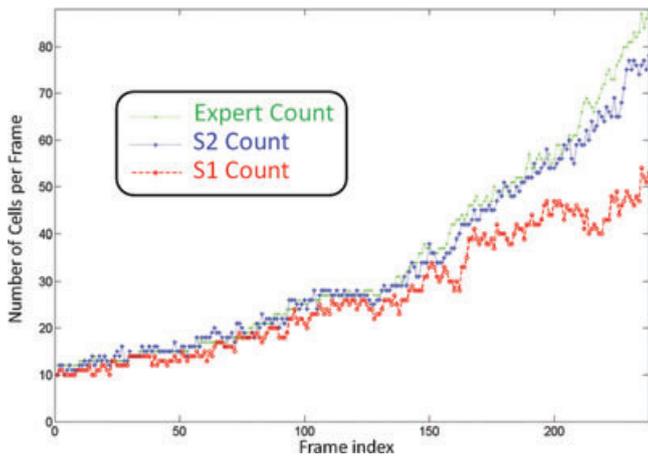


Fig. 10. Improved cell-counting accuracy after applying algorithms to segment adjacent cells in contact. NCF of the expert images (Expert Count), the segmentation with adjacent cell separation (S2 Count) and segmentation without adjacent cell separation (S1 Count) is plotted versus the frame index of the time-lapse image set.

(corresponding to the first 38 h of imaging). However, S1 begins to deviate much more from the reference data than S2 as the cell density and number of cell-to-cell contacts increase. For example, for the last 30 frames when many cells are adjacent, the difference between mean (ARI of S2) and mean (ARI of S1) is between 0.3 and 0.4. We conclude that on average the S2 method is 10% more accurate than the S1 method when looking at both ARI over the entire data set.

This improvement is also illustrated in Figure 10 where the number of correctly segmented cells by S2 is much closer to the reference set than the number of cells obtained by S1. The average deviation (over the entire FOV) in percent of number

of cells per frame between the result from S2 and the reference number of cells is 5.39% with a standard deviation of 4.6% ($n = 238$ images) whereas the average deviation between the result from S1 and the reference data set is 17.04% with a standard deviation of 10.9% ($n = 238$ images from a time-lapse image sequence).

The presented segmentation, in comparison with manual segmentation, has an average ARI of 0.96 (1 being a perfect match), with a standard deviation of 0.03, and an average difference of the two NCFs equals 5.39% with a standard deviation of 4.6% ($n = 238$ images from a time-lapse image sequence). In addition, when comparing the segmentation results with those based only on background pixel reconstruction, the incorporation of spatial and intensity models developed for dividing and nondividing cells resulted in a 10% improvement in the average accuracy and a 4-fold decrease in standard deviation based on the ARI metric. The decrease in the standard deviation suggests that in addition to being more accurate, the algorithm is also more precise for a given cell measurement. The same comparison based on the NCF metric resulted in a 11.65% improvement in the average accuracy and a 2.5 times smaller standard deviation.

Conclusions

We presented a segmentation procedure for Zernike phase contrast images of NIH 3T3 cells. The segmentation method is effective at separating cells when they are in contact. The novelty of the work lies in (a) taking advantage of live cell dynamics in the microscope FOV to achieve robust background image reconstruction, (b) incorporating biological models of dividing and nondividing cells into a computational model to deliver a more accurate and automated segmentation method capable of separating cells in contact and (c) validating the accuracy of segmentation results using manually created reference data consisting of a stack of 238 images. This segmentation technique has proven to be useful for fundamental live cell studies (Halter *et al.*, 2011), and is contributing to more accurate results for motion tracking of live cells as well. However, this approach requires that time-lapse images are acquired such that an accurate identification of the background pixel intensities can be made. This is facilitated by low initial cell seeding densities, migratory cells that expose background pixels and a uniform background that can be estimated when pixels are not exposed. In general, plotting the frequency of occurrence of background pixels in the time series stack, as shown in Figure 3(c), can indicate when the background reconstruction approach is likely to succeed. Background reconstruction will be most accurate when all or nearly all of the background pixels are exposed during the time-lapse acquisition. Our future work will focus on background reconstruction for microscopy images with a high cell density in the FOV, and on extensions of the biological models to other cell lines. The current algorithm

is segmenting only single frame. We plan to improve the algorithm in the future to include the analyses of previous and next frames which adds more information and help improving the accuracy of the current segmentation technique.

Disclaimer

No approval or endorsement of any commercial product by NIST is intended or implied. Certain commercial software are identified in this report to facilitate better understanding. Such identification does not imply recommendations or endorsement by NIST nor does it imply the software identified are necessarily the best available for the purpose.

References

- Bajcsy, P., Chalfoun, J. & Brady, M. (2012) Toward a Recommendation System for Image Similarity Metrics. *Proceedings of the 2nd IASTED International Symposia Imaging and Signal Processing in Health Care and Technology (ISPHT 2012)*, 94–100. Baltimore, MD. doi: 10.2316/P.2012.771-014.
- Bise, R., Li, K. & Eom, S. (2009) Reliably Tracking Partially Overlapping Neural Stem Cells in Dic Microscopy Image Sequences. *Proceedings of the MICCAI Workshop on Optical Tissue Image Analysis in Microscopy, Histopathology and Endoscopy*, 67–77, Imperial College, London.
- Bradhurst, C.J., Boles, W. & Xiao, Y. (2008) Segmentation of Bone Marrow Stromal Cells in Phase Contrast Microscopy Images. *Proceedings of the Image and Vision Computing New Zealand. 23rd International Conference*, 1–6. IEEE.
- Canny, J. (1986) A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8 (6) (November): 679–698. doi: 10.1109/TPAMI.1986.4767851.
- Debeir, O., Adanja, I. & Warzee, N. (2008) Phase Contrast Image Segmentation by Weak Watershed Transform Assembly. *Biomedical Imaging: From Nano to Macro, 2008. ISBT 2008. 5th IEEE International Symposium*, Paris, France, 724–727.
- Dima, A.A., Elliott, J.T., Filliben, J.J., et al. (2011) Comparison of segmentation algorithms for fluorescence microscopy images of cells. *Cytometry A* 79(7), 545–559. doi: 10.1002/cyto.a.21079.
- Dzyubachyk, O., van Cappellen, W., Essers, J., Niessen, W. & Meijering, E. (2009) Energy Minimization Methods for Cell Motion Correction and Intracellular Analysis in Live-cell Fluorescence Microscopy. *Proceedings of the 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro Boston, MA*, (June): 1127–1130. doi:10.1109/ISBI.2009.5193255.
- Ersoy, I. & Palaniappan, K. (2008) Multi-feature Contour Evolution for Automatic Live Cell Segmentation in Time Lapse Imagery. *Proceedings for the Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE* 1, 371–374. IEEE, Vancouver, Canada.
- Gonzalez, R.C. & Woods, R.E. (2008) *Digital Image Processing*. Pearson, Upper Saddle River, NJ. doi: 013168728X.
- Halter, M., Almeida, J.L., Tona, A., Kenneth, D., Plant, A.L. & Elliott, J.T. (2009) A mechanistically relevant cytotoxicity assay based on the detection of cellular green fluorescent protein. *Culture* 7(4), 1–9.
- Halter, M., Sisan, D.R., Chalfoun, J., et al. (2011) Cell cycle dependent tn-c promoter activity determined by live cell imaging. *Cytometry A* 79(3), 192–202.
- Hubert, L. & Arable, P. (1985) Comparing partitions. *J. Classif.* 2(1), 193–218. doi: 10.1007/BF01908075.
- Maurer, C.R., Jr. & Qi, R. (2003) A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *Pattern Anal. Mach. Intell.* 25(2), 265–270.
- Kachouie, N.N., Fieguth, P. & Jervis, E. (2007) Stem-cell Localization: A Deconvolution Problem. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France*. (January), 5525–5528. doi: 10.1109/IEMBS.2007.4353597.
- Kang, W., Lai, W. & Meng, X. (2009) An adaptive background reconstruction algorithm based on inertial filtering. *Optoelectronics Lett.* 5 (6), 468–471. <http://www.springerlink.com/index/Y16057652322L681.pdf>.
- Kociolek, Marcin. Canyon Software (2009). http://www.eletel.p.lodz.pl/kociolek/Download/Canyon_v07.exe, accessed Dec 2011.
- Langenbach, K.J., Elliott, J.T., Tona, A., McDaniel, D. & Plant, A.L. (2006) Thin films of type 1 collagen for cell by cell analysis of morphology and tenascin-C promoter activity. *BMC Biotechnology*. 6, 6–14.
- Roerdink, J.B.T.M. & Meijster, A. (2001) The watershed transform: definitions, algorithms and parallelization strategies. *Fundamenta Informaticae* 41, 1–40.
- Sigal, A., Milo, R., Cohen, A., et al. (2006a) Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins. *Nat. Methods* 3(7), 525–531.
- Sigal, A., Milo, R., Cohen, A., et al. (2006b) Variability and memory of protein levels in human cells. *Nature* 444(7119), 643–646. doi: 10.1038/nature05316.
- Smolka, J. (2005) Watershed based region growing algorithm. *Annales UMCS Informatica AI* 3, 169–178.
- Tang, J., Jing, X. & He, D. (2008) Research on Background Reconstruction Based on Prediction and Filtering. *Proceedings of the Computer Science and Software Engineering, 2008 International Conference On* 1, 969–972. IEEE, Montreal, Canada.
- Tscherepanow, M., Zöllner, F., Hillebrand, M. & Kummert, F. (2008) Automatic segmentation of unstained living cells in bright-field microscope images. *Proceedings of the Advances in Mass Data Analysis of Images and Signals in Medicine, Biotechnology, Chemistry and Food Industry* 5108, 158–172, Leipzig, Germany.
- Vinh, N.X., Epps, J. & Bailey, J. (2009) Information Theoretic Measures for Clusterings Comparison: Is a Correction for Chance Necessary? *Proceedings of the 26th Annual International Conference on Machine Learning*, 1073–1080. ACM, Montreal, Canada.
- Wang, M., Zhou, X., Li, F., Huckins, J., King, R.W. & Wong, S.T.C. (2008) Novel cell segmentation and online svm for cell cycle phase identification in automated microscopy. *Bioinformatics (Oxford, England)* 24(1), 94–101. doi: 10.1093/bioinformatics/btm530.
- Williams, D. & Shah, M. (1992) A Fast Algorithm for Active Contours and Curvature Estimation. *CVGIP: Image Understanding*. 55(1), 14–26. doi:10.1016/1049-9660(92)90003-L.
- Zanella, C., Campana, M., Rizzi, B., et al. (2010) Cells Segmentation from 3-D Confocal Images of Early Zebrafish Embryogenesis. *Proceedings*

- of the *IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society* **19**(3), 770–781, Piscataway, NJ. doi: 10.1109/TIP.2009.2033629.
- Zijdenbos, A.P., Dawant, B.M., Margolin, R.A. & Palmer, A.C. (1994) Morphometric Analysis of White Matter Lesions in MR Images: Method and Validation. *Proceedings of the IEEE Transactions on Medical Imaging* **13**(4), 716–724. doi: 10.1109/42.363096.
- Zou, K.H., Warfield, S.K., Bharatha, A., Tempany, C.M.C., Kaus, M.R., Haker, S.J., Wells Iii, W.M. & Jolesz, F.A. (2006) Statistical validation of image segmentation quality based on a spatial overlap index. *NIH Public Access* **11**(2), 178–189.