Journal of Imaging Science and Technology® 52(3): 030509–030509-23, 2008.

High-Throughput Identification and Classification Algorithm for Leukemia Population Statistics

Brinda Prasad and Wael Badawy
Department of Electrical and Computer Engineering, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4
E-mail: bprasad@smarttech.com

Abstract. *Early detection of leukemia and reduced risk to human health can result from interdisciplinary integration of image analysis with clinical experimental results. Image analysis relies on efficient and reliable processing algorithms to make quantitative judgments on image data. This article presents the design and implementation of an efficient and high-throughput leukemia cell count and cluster classification algorithm to automatically quantify leukemia population statistics in the field of view. The algorithm is divided into two stages: (1) the cell identification stage and (2) the cell classification and inspection stage. The cell identification stage accurately segments background and noise from foreground pixels. A boundary box is generated enclosing the foreground pixels identifying all isolated cells and cell clusters. The cell classification and inspection stage uses one-dimensional intensity profiles that behave as signature plots to segregate isolated cells from cell clusters and evaluate total count within each cluster. The designed algorithm is tested with a variety of leukemia cell images that vary in image acquisition conditions, image sizes, cell sizes, intensity distributions, and image quality. The proposed algorithm demonstrates good potential in processing both ideal and nonideal images with an average accuracy of 91% and average processing time of* 3 s. *The performance of the proposed algorithm in comparison to recently published algorithms and commercial image analysis tool further ascertains its robustness.* © *2008 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.ImagingSci.Technol.(2008)52:3(030509)]

INTRODUCTION

Leukemia is a bone marrow disorder that arises when white blood cells (WBCs) abnormally begin to replicate themselves.[1] Leukemia cells do not function normally, fail to fight infection as they should, have increased propensity to survive conditions of growth and do not die at the same rate as other WBCs.[2] Quantitative analysis of WBC count could assist in monitoring the growth rate of leukemia cells to drug actions or radiation treatments like chemotherapy.[3] WBC counts that continue to rise or fall to abnormal levels indicate medical emergency, and conversely, counts that return to normal indicate improvement.[4]

Digital image processing has the advantage of rapid processing and quantitative extraction of features from regions of an image, a task that humans can only estimate using qualitative means. Developing algorithms to analyze leukemia cell images is challenging owing to the variability of features and characteristics of the cell samples. In addition, the image acquisition conditions dictate the image quality that could immensely affect the accuracy and sensitivity of the algorithm. This article presents the design and implementation of a high-throughput cell count and cluster classification algorithm to automatically quantify the distribution of leukemia cells in a field of view. In order to locate each cell and accurately classify clusters, correct segmentation of not only solitary cells but also complex cell clusters needs to be performed because ignoring clusters affects the differential count.[5] Although the presence of cell cluster constellations are rare in healthy peripheral blood, they are abundant in pathological conditions with high leukocyte concentrations, such as leukemia.[5] Typically, all cultures of leukemia cells contain a fraction of asymmetric pear-shaped cells with a single cytoplasmic protrusion.[6] These protrusions vary in size from about one-third to one-eighth of the diameter of the main cell body. Cytoplasmic material is asymmetrically distributed with most particulate organelles at the side of the protrusions.[6]

Leukemia cells tend to adhere to one another even at very low densities, at which stage small aggregates predominate, consisting of two to four cells characteristically joined together by their protrusions. As the cell density increases, larger cell clusters of no definable organization appear in increasing numbers. Human operators may not be able to accurately detect the correct count in the presence of such clusters or clumps. Separating individual cells from a cell clump can be nontrivial since they are closely bound to each other and may not always have defined edges, as shown in Figure 1. Ignoring cell clusters in the population analysis of cell samples could affect the differential count that may be propagated to feature calculations and eventually lead to inaccurate diagnosis. Hence, analysis and classification of cell clumps is a vital step to ensure overall success of the obtained population statistics. The purpose of the designed algorithm is also to achieve maximum accuracy with a wide variety of leukemia cell samples. With this intention in mind, the proposed algorithm is targeted toward leukemia cells, both in bright field [Figs. 1(a)–1(c)] and fluorescently stained images [Fig. 1(d)] in dark fields that vary in image and cell sizes, noise levels, intensity distributions, and image quality.

The identification of most cells in microscopic cytology is traditionally performed manually by experienced labora-

Received Jul. 14, 2007; accepted for publication Feb. 21, 2008; published online Jun. 12, 2008.
1062-3701/2008/52(3)/030509/23/$20.00.

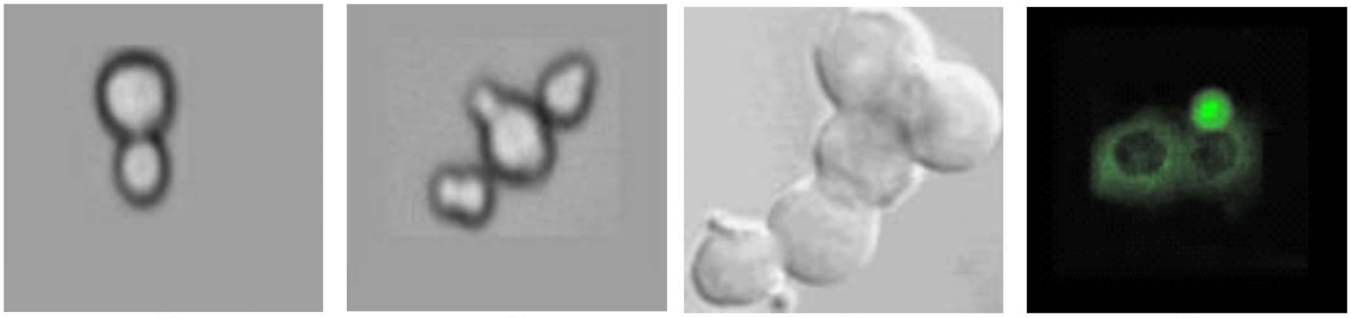


Figure 1. Cell clusters with joined edges; (a) cluster of two cells, (b) cluster of three cells, (c) cluster of group of cells, and (d) cluster with fluorescent stain (green colored).

tory technicians. This approach is tedious and time consuming, especially when analyzing images with a large number of cells. A semiautomatic system was developed to estimate proliferating cell nuclear antigen (PCNA) expression in paraffin sections of breast carcinomas.[7] Selected tumor cells from marked areas were manually graded by the user. Automated cell counting algorithms in the past have been based on characteristic feature thresholding or feature clustering,[8,9] edge detection techniques,[10] region extraction,[11–14] or a combination of the these methods.[15–17] The methods described in these algorithms targeted images on peripheral blood cells that are straightforward because the cells were sparsely distributed. Wermser et al.[18] proposed a fast and simple segmentation scheme based on hierarchical thresholding for automatic differential blood count and the decomposition of blood cells into the nucleus, plasma, erythrocytes, and thrombocyte. Poon et al.[19] proposed an algorithm to detect the presence of just-touching cells. However, this algorithm separates pairs of cells but cannot handle more complex cell cluster constellations. A cluster division algorithm for dividing paths across clusters was proposed by Lockett and Herman.[20] A dividing path was defined as the path that possessed an average gradient per pixel greater than all other possible paths. This approach fails in analyzing thicker specimens (>50 μm), where greater overlap of cells is expected. Moreover, in our case, the fluorescently stained test images being processed contain cell clusters that are sometimes stained only at the edges (only cytoplasm, not the nuclei). A dividing path to separate such stained clusters would not be feasible. A method to efficiently segment nuclear clusters, based on watershed lines, was demonstrated by Malpica et al.[21] An incorrect choice of the image transformation would lead to incorrect segmentation, and if a watershed transform was performed using all local minima of the gradient image, an oversegmented image resulted, in which all the contours between two minima is present, as shown in Figure 2. Sjöstrom et al.[22]proposed a cell counting system based on an artificial neural network (ANN) that used an intelligent filter to scan regions of interest in the input image to reject debris and other objects classified as noncells. However, these authors report a limited degree of accuracy and high computational overhead.

Image processing techniques based on processing histograms to find regions of interest in cancer cell nuclei were suggested by Loukas and Vojnovic.[23] This technique was unable to effectively segment the regions of interest of cell images with histological noise and nonuniform background. A blood cell counting and segmenting technique based on the feature of the cell's logical and morphological information was proposed by Ma et al.[24] These authors reported that the counting based on their technique is not very accurate. A two-stage tumor cell identification strategy to automatically segment tumor cells using local adaptive thresholding and dynamic water immersion techniques was proposed.[25] These authors report a number of false identifications that are obtained due to histological noise in the images. An automated differential counting of blood leukocytes based on automatic cell location and image processing for the preclassification of blood cells using artificial neural networks was proposed by Swolin et al.[26] The objective of the paper was to use neural networks to classify and quantify normal cell classes versus abnormal cells based on cell morphology. Most closely packed cells are not convex, thus morphologic operators cannot be used to separate touching cells. Nilsson and Heyden[27] proposed an algorithm to segment complex cell clusters for bone marrow samples by solving a combinatorial optimization problem. The algorithm, however, imposes a maximum limit on the diameter of cells. A new variational formulation for geometric active contours that forces the level set function to be close to a signed distance function was proposed by Li et al.[28] A cell count based on motion based segmentation algorithm was suggested by Prasad et al.[29] This algorithm uses motion as a cue to use an adaptive mesh to identify cell boundaries. No efforts toward cluster analysis were discussed with the motion capture confined to two to seven cells only.[29]

A robust rule-based approach for the splitting of binary clumps that are formed by objects of diverse shapes and sizes was proposed by Kumar et al.[30] This method was based on splitting a binary clump into two smaller clumps until no more split lines were detected. The method began with the detection of concavity pixels in a clump. The boundary arcs and the convex hull segments of the clump were initially obtained such that the pixel with the largest perpendicular distance from the corresponding convex hull was selected as the concavity pixel. Split lines were selected by joining all possible pairs of concavity pixels. A cost function was also introduced for determining the best split line from the set of

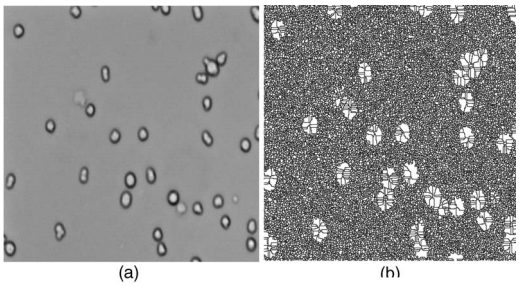


Figure 2. Leukemia *FMC44* cell type acquired in bright field oversegmented using watershed algorithm.

candidate split lines. The authors do report that the clump splitting technique accurately splits clusters that have only one major concavity region. This method closely resembles a recently proposed convex hull based technique to trace and classify isolated cells and cell cluster.[31] The percentage of the actual area of the cell against the area of the convex hull was calculated and corresponding height profiles were analyzed to determine a single cell from a cell cluster. An estimation of cultured rat B104 neuroblastoma cell cluster count by sorting cells based on a representative cell was suggested by Wu et al.[32] The authors reported that their proposed algorithm was based on the assumption that cells within the same sample are alike. To quantitatively address a variety of biological questions, a free, open-source system designed for flexible high-throughput cell image analysis, Cell Profiler, was recently developed.[33] Commercially available image analysis tools conventionally used in some laboratories also provide cell counting feature.[34–36] However, the tools are invariably expensive and highly specific to the image being processed.

The rest of this article is organized as follows. Section II provides the proposed algorithmic description. Section III illustrates and discusses the results of the algorithm on a variety of leukemia cell samples. The performance of the proposed technique is analyzed by comparison with other proposed techniques in Section IV. Section V presents conclusions.

PROPOSED ALGORITHMIC DESCRIPTION

The challenge facing imaging researchers lies in developing smart image processing algorithms that can effectively and efficiently extract and quantify regions of biological importance. The challenge becomes particularly daunting when the designed algorithm needs to be efficiently applicable to a variety of images of different cell sizes, intensity distributions, and noise levels. In order to maintain optimum computational speed, the processing complexity that the algorithm can handle is also restricted. This section discusses the proposed cell identification and classification algorithm that is designed to identify and classify each cell as a single cell or as a cell cluster and to inspect and analyze cell count within each cluster. To this effect the proposed algorithm is divided into two main stages:

- Stage 1. Cell identification. In this stage, the cell image is accurately segmented to identify isolated cells and cell clusters from the background pixels and noise.
- Stage 2. Cell classification and inspection. An intensity summation plot of all the identified cells is then analyzed. These plots behave as signature profiles to segregate isolated cells from cell clusters and determine the total number of cells within the cluster.

The population statistics, along with other relevant statistics, are then quantified and displayed on the graphical user interface (GUI). The image acquisition details, along with processing steps used by the proposed algorithm, will be discussed in the following subsections.

Image Acquisition

The cell samples are placed on the Olympus microscope that supports bright field and fluorescent modes.[37] Olympus is a powerful universal microscope designed for laboratory investigations that embodies the modular construction principle of binocular and trinocular microscopes and provides a magnification range of 10× to 1250×. The in-built illuminator is based on a 30 W halogen bulb light source. It is feasible to attach phase-contrast or dark-field condensers, a fluorescence box, polarizers, nonimmersion and water-immersion lenses. The microscope is interfaced to the Olympus digital camera (model number DP25) and the PC using a fire-wire connection (IEEE 1394). The DP25 is a 5 megapixel 2/3 in. interlaced CCD digital color camera with a C-mount head having a pixel resolution of 3.4 μm with four image resolutions: 4080×3072, 2040×1536, 1360×1024, and 680×512.[37] The Olympus microscope provides a rapid changeover from the bright-field mode to fluorescence dark-field mode as desired. A dark-field mode provides the highest contrast when one works with nontransparent samples. A bright-field mode provides a sharp image in all feasible magnifications. The image acquisition conditions vary for each of the test images with the acquired image size with horizontal and vertical resolutions ranging from 256×256 to 1024×2048. Elaborate details highlighting cell type, experimental, and the acquisition details of test images are as follows:

(1) Image name: 3e10, FMC44. Leukemia cell lines acquired in bright field with an Olympus BX-51 microscope and captured using Image Pro software (1000×1000 resolution); magnification (10× objective and 10× eyepiece).
(2) Image name: 3t39brev. Leukemia cell lines acquired in bright field with an Olympus BX-51 microscope and captured using Image Pro software (256×256 resolution); magnification (40× objective and 10× eye-piece).
(3) Image name: CD44i. Leukemia cell marker acquired in bright field with an Olympus BX-51 microscope and captured using Image Pro software (256×256 resolution); magnification (40× objective and 10× eye-piece).
(4) Image name: *MMB1/MMB3*. Multiple myeloma

acquired in bright field with an Olympus BX-51 microscope and captured using Image Pro Software (256×256 resolution); magnification (20× objective and 10× eye-piece).

(5) Image name: *MMB1-red/MMB3-red*. Multiple myeloma acquired with an Olympus BX-51 microscope in fluorescence mode and captured using Image Pro software (256×256 resolution); magnification (10× objective and 10× eye-piece); Cells stained using Alexa red dye; excitation wavelength, 594 nm; emission wavelength, 617 nm.

(6) Image name: *SG28RFLt2-green*. Transfectants acquired in dark field with an Olympus BX-51 microscope in fluorscent mode captured using Image Pro software (256×256 resolution) magnification (10× objective and 10× eye-piece). Stained using Enhanced Green Fluorescent Protein (EGFP).

(7) Image name: *CD44i-red*. Cell marker/cell line acquired with an Olympus BX-51 microscope in fluorescence mode and captured using Image Pro software (256×256 resolution); magnification (40× objective and 10× eye-piece); Cells stained using Alexa red dye; excitation wavelength, 594 nm; emission wavelength, 617 nm.

(8) Image name: *3t39brev-mixed*. Leukemia cell line acquired with an Olympus BX-51 microscope in fluorescence mode and captured using Image Pro Software (256×256 resolution); magnification (40× objective and 10× eye-piece); Cells stained using Alexa red+green dye.

(9) Image name: Hemo1/Hemo2/Hemo3. A hematopoietic cell line derived from Burketts lymphoma (*Daudi*) was purchased from ATCC (American Type Culture Collection, Manassas, VA, USA). *Daudi* is a human B (lymphoblast) cell line originally derived from peripheral blood of a Burketts lymphoma patient. The cells are nonadherent cells that multiply in suspension cultures and reach stationary phase in approximately three to four days. The cell line and culture medium (RPMI 1640) were supplied by Cross Cancer Institute (CCI, Edmonton, Canada). The cell line was maintained in the RPMI 1640 medium (Gibco, Invitrogen, Canada) and supplemented with 10% fetal bovine serum (RPMI 1640 11875-093. Fetal Bovine Serum, Quantified 10437- 028) in an incubator where the cells were kept at 5% CO_2, 95% filtered breathing grade air at 37°. The CO_2 level inside the incubator was checked with a Fyrite one to two times every month. The cultured *Daudi* cells were passed every three to four days. All cell manipulations were performed at room temperature of approximately 20° in a class-A bio-safety cabinet (NU-425, Nuaire, Inc., Plymouth, MN, USA). All images were taken through a 10× objective (additional 10× with the eye-piece) on an Olympus BX-51 microscope and captured using Image Pro Software (1024×2048 resolution). The percentage of live/dead *Daudi* cells were determined using the Eosin dye exclusion method (12 ml sample and 12 ml Eosin dye) and counting via a hemocytometer. Dead cells absorb the dye and appear as red dots under the microscope, whereas healthy live cells appear as rings.

Cell Identification

The algorithm begins with a preprocessing stage intended to reduce noise to reduce propagation to subsequent steps. The noise level in images may vary, which could result from poor lighting or poor focusing conditions during acquisition or from dust particles on the microscope slide. A nonlinear 3×3 median filter was used to eliminate the noisy pixels from the input image. This type of spatial filter reduces noisy pixels that appear as salt-and-pepper noise, but preserves the high frequency pixels causing minimal changes to noise-free parts of the image.[38]

The proposed algorithm is then subjected to a red-green-blue (RGB) color channel extraction. The RGB color channels are stripped from the cell image and the component video channel is generated by separating a luminance component (Y) from RGB channels. The algorithm then goes through a series of processing steps that includes edge detection, thresholding, recursive segmentation, and boundary extraction steps. Elaborate discussion regarding the cell identification stage is presented in our earlier publications.[39,40] Figures 3(a)–3(d) illustrate the processing outcomes in stages for the cell identification stage using leukemia cell image with 17G10 cell type acquired under bright field.[39] The generated boundary represented by a white boundary box as shown in Fig. 3(d) encloses all the identified cells in the image.

Cell Classification and Inspection

The cell classification and inspection stage in stage 2 of the proposed algorithm uses Fig. 3(d) from stage 1 as an input to classify cells as either single cell or cell clusters. Cell clusters are further inspected to obtain complete population statistics. The processing steps involved in the classification and inspection stage are elaborated in the subsequent sections.

(1) Diameter check. The first step in the cell classification stage is to separate single cells from cell clusters. This is done by computing the diameter of all identified cells from the coordinates of the generated cell boundaries, as shown in Figure 4. The average diameter is estimated such that all cells with a diameter less than the average diameter are classified as single cells represented by white boundaries, as shown in Figure 5. However, 10%–15% of the single cells that are larger in size are incorrectly classified as cell clusters represented with black boundaries, as in Fig. 5.

Figure 6 shows the diameter computed from boundary box dimensions for all isolated cells from processed test image 1, test image 2, and test image 3. It can be seen that a large variation in cell sizes

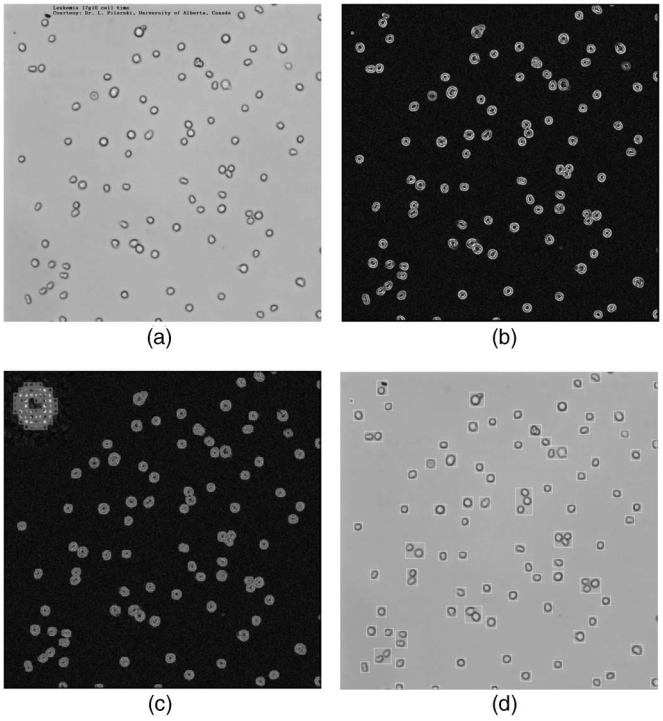


Figure 3. Processed results from Stage 1—identification stage of the proposed algorithm.[39]

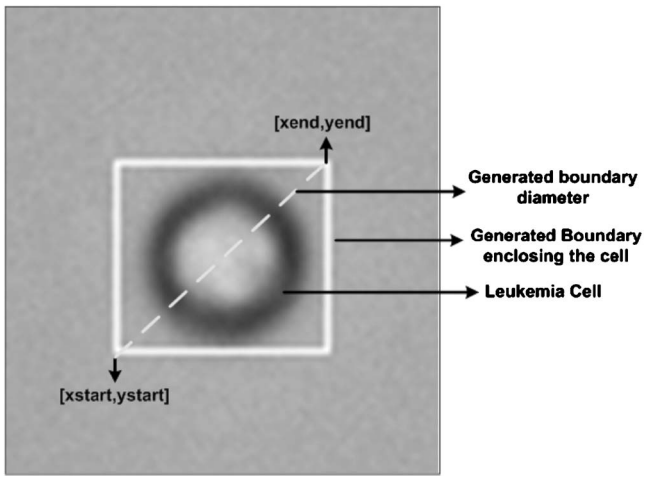


Figure 4. Diameter computed from boundary coordinates.

(in pixels) could be observed within the same sample at a given time. This re-emphasizes the fact that the classification of cells based only on the size as used by Wu et al.[32] may not be accurate in cell classification.

(2) Median check. At this point a 1D intensity summation plot both row-wise and columnwise is plotted for all identified cells. A typical 1D plot for a single cell versus cell cluster is shown in Figures 7(a) and 7(b). The y-axis (x-axis) of the 1D plot corresponds to the total sum of intensities in a column (row). The total number of columns (rows) is

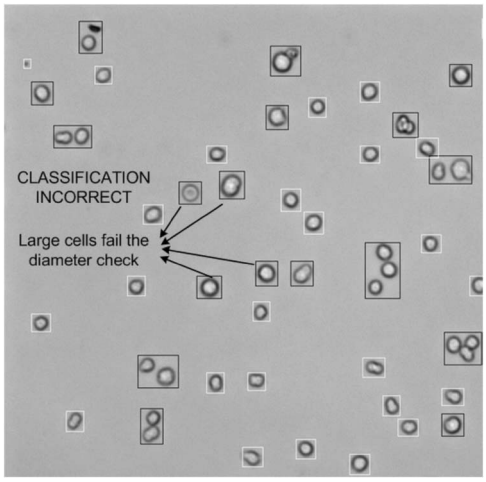


Figure 5. Processed results after diameter check. Single cells with large diameter incorrectly classified as a cell cluster.

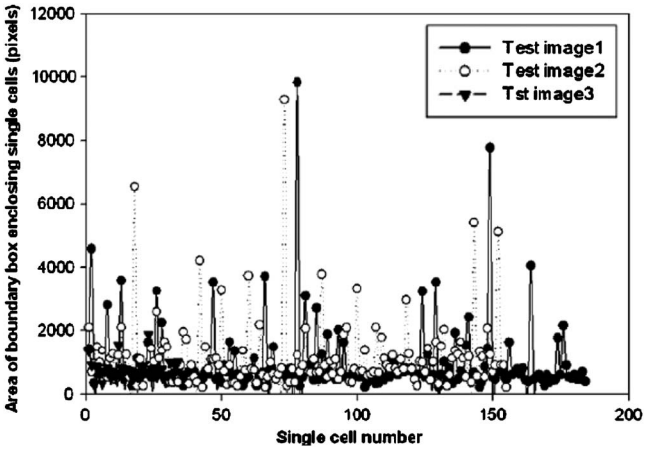


Figure 6. Size distribution of boundary box enclosing single cells on daudi test cell samples on a hemocytometer.

determined from the generated boundary box dimensions. For example, if a cell with a generated boundary box dimension is $[x_{start}, y_{start}]=[20,40]$ and $[x_{end}, y_{end}]=[60,80]$, the summation plot would have 40 (80−40) columns and 40 (60−20) rows.

Base –C (Base –R) refers to the count of intensity summation values for each column (row) above an optimum value for a particular cell or cluster. This optimum value varies with each image and is adaptively computed by the proposed algorithm. Base – C – zero (Base –R – zero) refers to the count of intensity summation values below an optimum value for a particular cell or cluster. For example, if the optimum value is determined as 1000 for an image, Base – C would determine count of all the intensity summation for every column (within the boundary dimension) above 1000 and Base –C –zero would represent count of all the intensity summation for every column below 1000. Base – C – total (Base – R – total) is the sum of

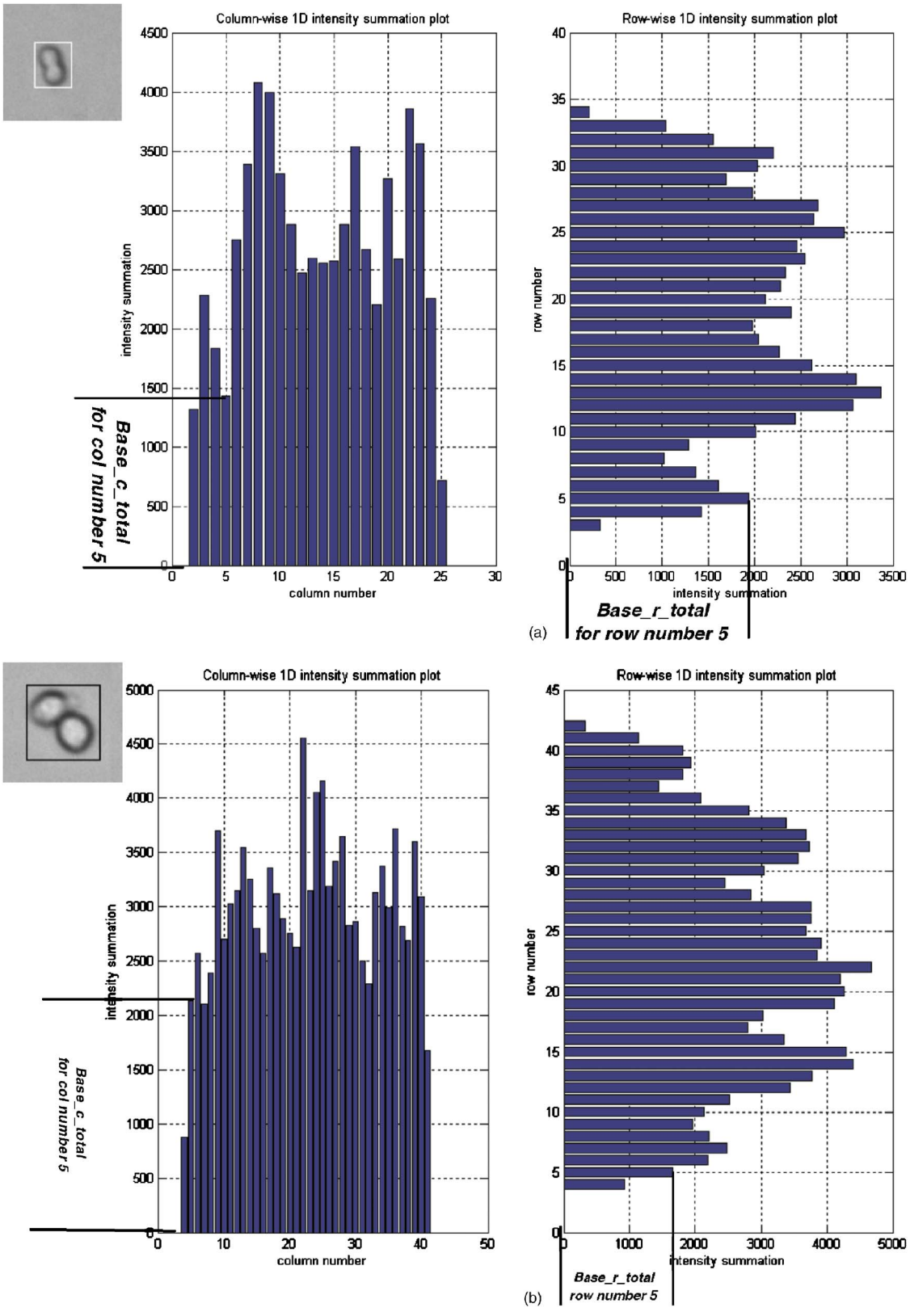


Figure 7. 1D intensity profiles of single cell and cell cluster.

Base – C (Base – R) and Base – C – zero (Base – R – zero). All the previous attributes are computed for all the cells with their generated cell boundaries. The attributes are abbreviated as BCT (Base –C –total), BRT (Base –R –total), BCZ (Base –C –zero), and BRZ (Base – R – zero), respectively. The incorrectly classified cells from the diameter check are subjected to another test called the "median check". A median value from the list of all the computed BCT (BRT) is determined. All cells with BCT (BRT) lesser than the median value are declared as isolated cells. All other cells are declared as cell clusters. Figure 8 shows the processed results after a median check. As seen in Fig. 8, all the single isolated cells that were initially incorrectly classified (Fig. 5) are now accurately classified as single cells. Thus, the cells are now classified as single cells represented by white boundaries, and cell clusters represented by black boundaries, as shown in Fig. 8. Preliminary simulations using standard blob analysis algorithms produce less than accurate results on sample test images. Figure 9 shows the processed results depicting centroids (green dots) on identified cells that do not accurately represent the cell's centroid. This prompted us to go with the intensity profile analysis as described in this article.

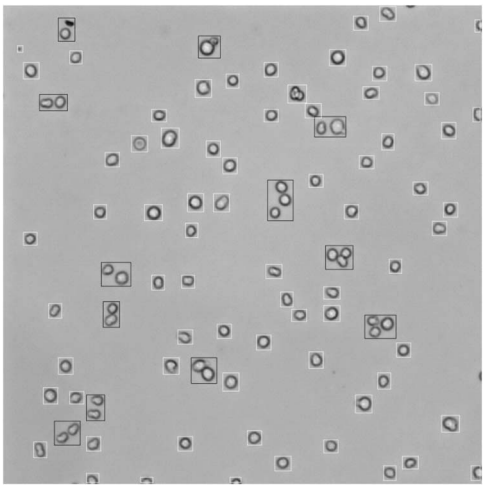
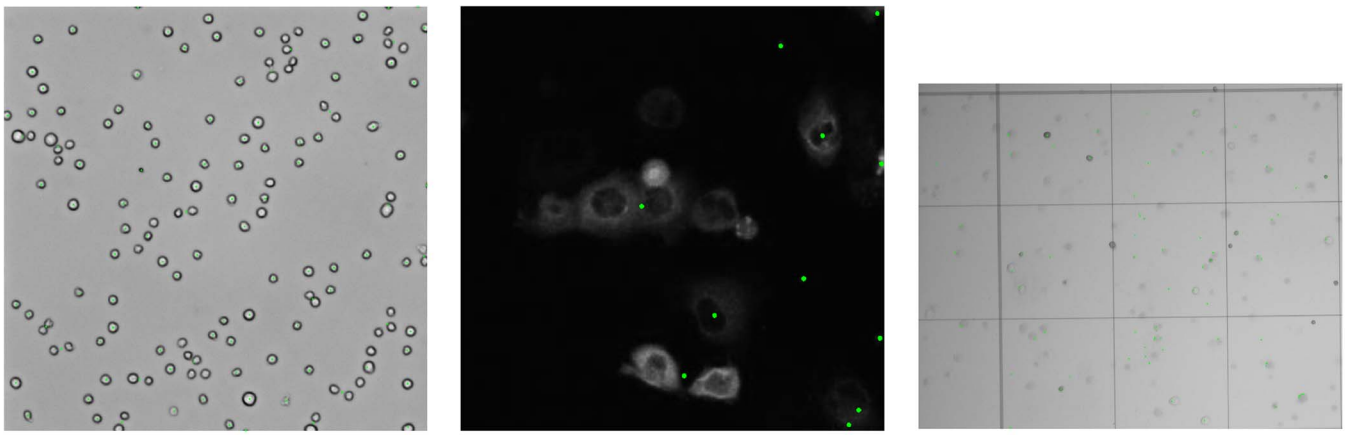


Figure 8. Processed results after median check. Single cells with large diameter accurately classified.

(3) Closely spaced cell separation. The median check from the previous step categorizes cells as single cells and cell clusters. However, along with cell clusters there exists another category of cells called the closely spaced cells. These cells contain multiple cells that are very close to each other. The recursive segmentation process in the identification stage failed to identify closely spaced cells because of close proximity to two or more cells. Although further recursive image segmentation (Stage 1–Step 2)[39] to separate all cells is an obvious solution, there is an exponential increase in computational cost, which has direct impact on the processing time. It is counterproductive to the goal of developing a fast and robust algorithm. The closely spaced cells could either be two or more closely spaced clusters or single cells or a combination of the two, as shown in Figures 10(a)–10(c).

Repeated iterations to separate the closely spaced cells is performed such that the separ04806prerated cells either fall in the single cell category or a cell cluster category. Hence, this step involves separation of closely spaced cells and reclassification of separated cells as a single cell or cell cluster. The 1D intensity profiles of closely spaced cells are inspected such that the spacing between cells are observed as visible gaps on the profile, as seen in Figure 11. A dividing line corresponding to the position of the gap can assist in the separation of such cells and cell groups. This process of separating cells with gaps is repeatedly performed until



(a) (b) (c)

Figure 9. Processed results from standard blob analysis algorithm depicting centroids of identified cells are depicted as green dots.

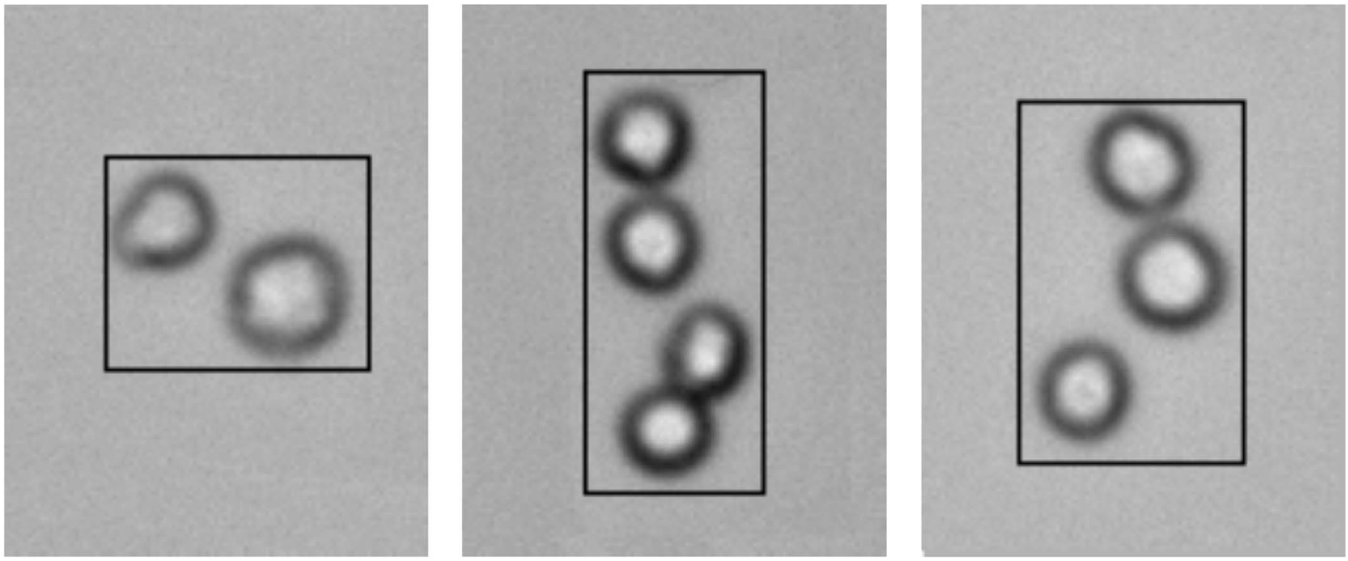


Figure 10. Examples of different combinations of closely spaced cells.

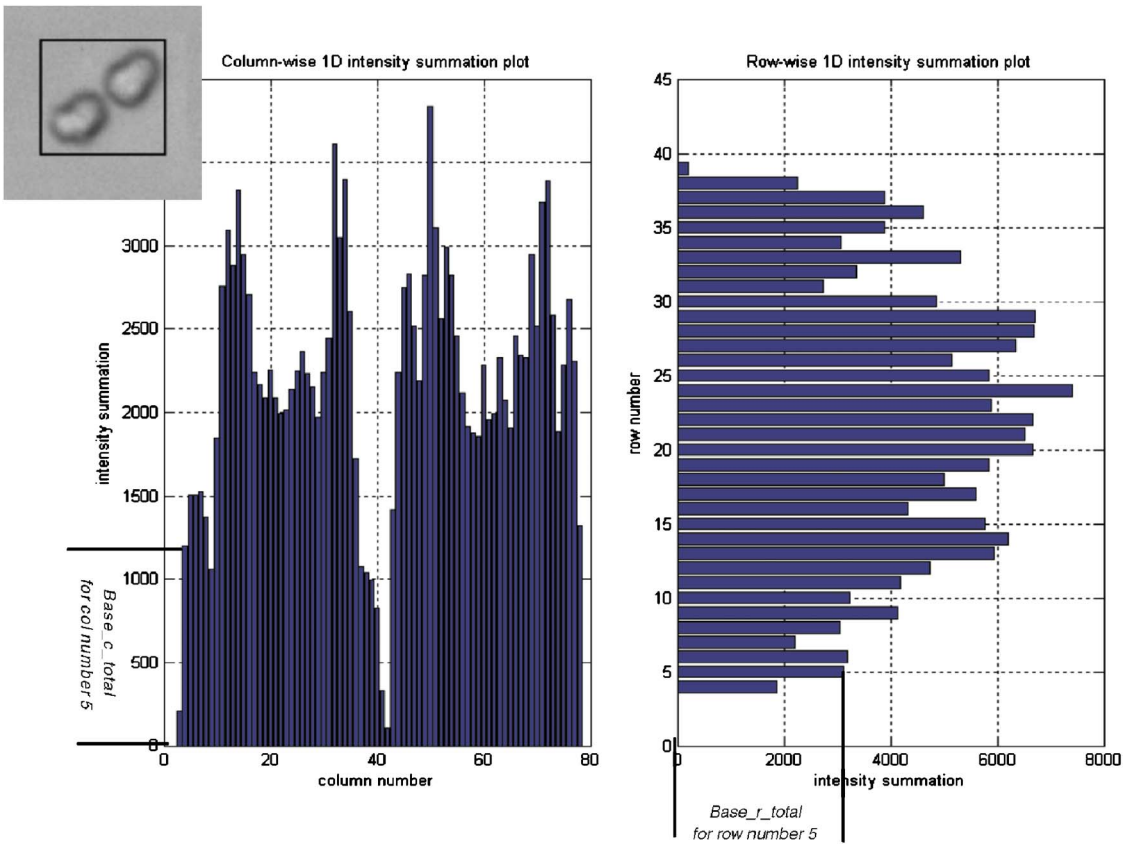


Figure 11. 1D intensity profiles of closely spaced cells with gaps in them to enable separation.

no more gaps exist in the 1D profile. This is done to ensure separation of cells with more than one gap. Intensity profiles and the corresponding attributes of the two newly divided groups are reestablished. BCT and BRT values are examined for all the newly divided cells to reclassify them as either a single cell or cell cluster. Hence, after this step only two groups of cells exist—isolated single cells and cell clusters. Figure 12 illustrates a step by step approach towards separating and classifying a closely spaced cell. As seen in Fig. 12 dividing lines corresponding to the gap locations help in separating a closely spaced cell into three isolated cells and one cluster. Figure 13 shows the processed results of the designed algorithm before and after Step 3. The single cells are represented with white boundary

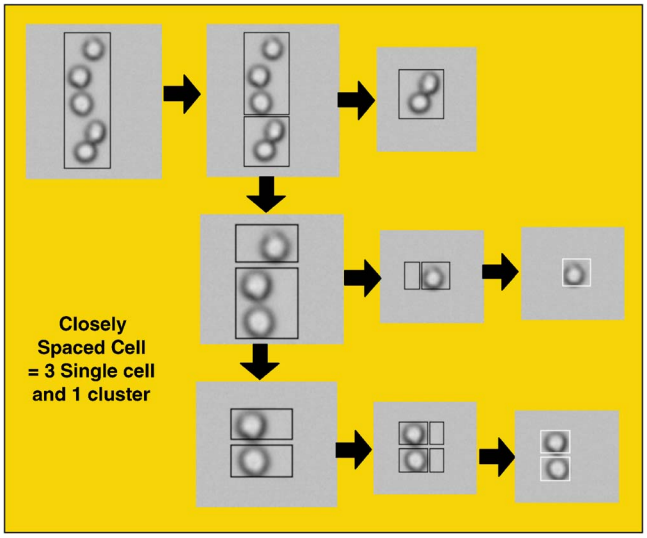


Figure 12. Step by step approach of separating closely spaced cells from their intensity profiles.

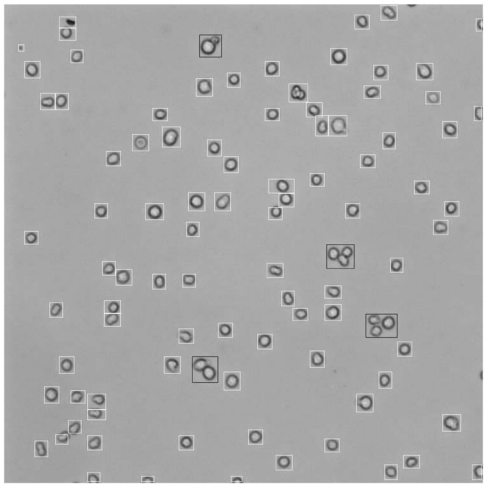


Figure 13. Processed results before and after closely spaced cell separation.

boxes and the cell clusters with black boundary boxes.

(4) Cluster analysis. After the cell clusters are accurately identified and classified, the next stage in the algorithm is to speculate every cell cluster to identify and determine the total number of cells that make up each cell cluster. The following tests are done to determine the cell count within a cell cluster:

• Test 1. At this point, all the isolated cell (single cells) locations have been tagged and will not participate in further processing. The related attributes to single cells are also collected for final populations. The first test to analyze clusters was to reuse their computed 1D intensity summation plots. It was observed that as the number of cells within a cluster increased by one there was a proportional increase in values of BCT–BCZ or BRT–BRZ or both by approximately 1.5–2.0 times. Figure 14 illustrates the 1D summation plots for a cluster with three cells and with two cells. It can be observed that the BCT–BCZ for a cluster of three cells is 1.55 times greater than a cluster with two cells.

• Test 2. In order to further validate the results from Test 1, a second level of testing is implemented to analyze cell count within clusters. The method involves applying dilation operation to the thresholded image obtained from Stage 1–Step 3[39] of the proposed algorithm. The dilation operator takes two pieces of data as inputs. The first is the image that is to be dilated. The second is a (usually small) set of coordinate points known as a structuring element (also known as a kernel). It is this structuring element that determines the precise effect of the dilation on the input image. Though dilation is applied to the entire binary image, the focus of attention is toward identified clusters that are enclosed by generated boundaries from the previous steps. Hence, the effect of dilation in sometimes highlighting noisy pixels can easily be ignored since they do not affect further processing steps. In our case, the input image is the thresholded binary image and the structuring element used is as follows:

$$\text{Mask-Kernel} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}. \quad (1)$$

Though dilation is applied to the entire binary image, the focus of attention is toward the identified clusters that are enclosed by generated boundaries from the previous steps. Hence, the effect of dilation that highlights noisy pixels can easily be ignored and they do not affect further processing steps. Once the dilated image is obtained, the boundary coordinates across the clusters are used to obtain the left, right, top, and bottom diagonals, as shown in Figure 15. Knowing the boundary coordinates, all the previous diagonals could easily be obtained using the equation of a straight line. Cumulative counts of foreground pixels (pixels with gray level 255) on the dilated image are computed across each pixel on the left, right, top, and bottom diagonals (cumulative count coined as CCFP for simplicity). As the cell count within the cluster increases, the length of diagonals and the pixels to analyzed across each diagonal correspondingly increases. Intuitively, CCFP across the diagonals also proportionally increases by 1.5–2.0 times with every increase in the cell count within the cluster. Table I summarizes the conditions used by the proposed algorithm to conclusively determine the cell

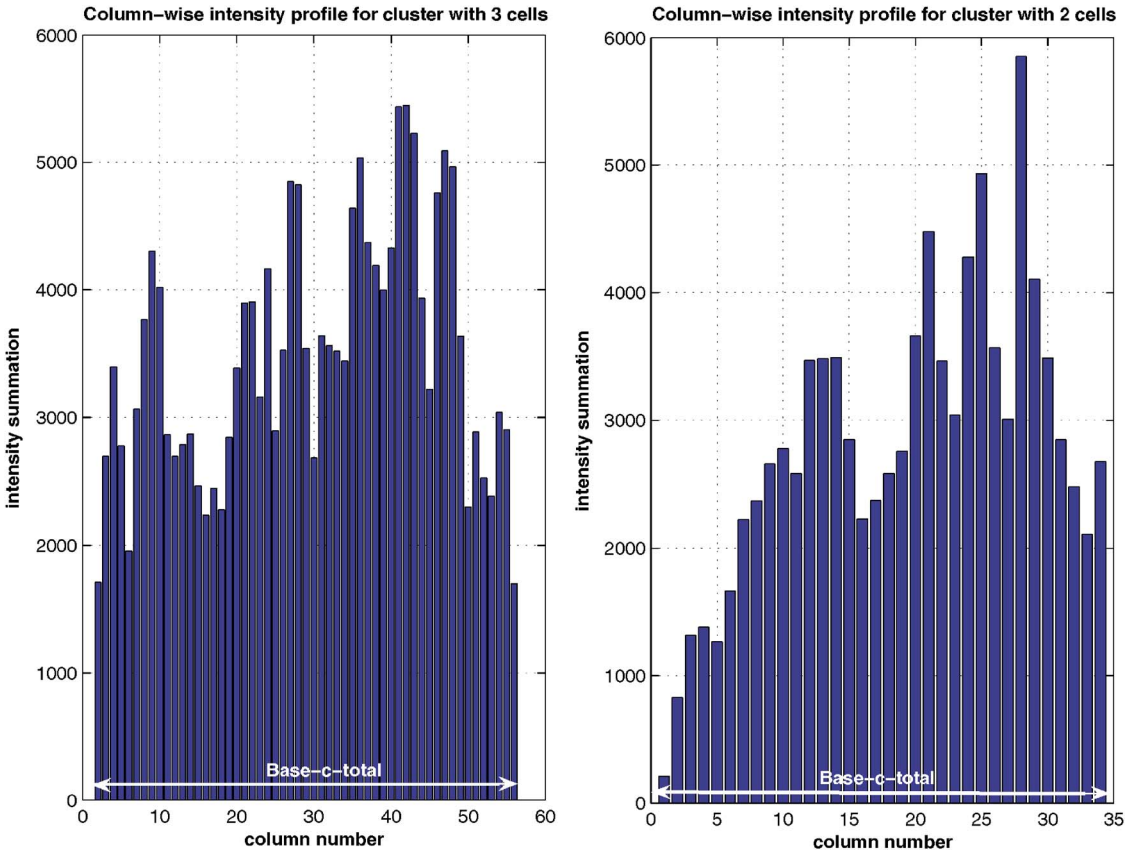


Figure 14. Comparisons of profiles with cluster of two cells with cluster of three cells.

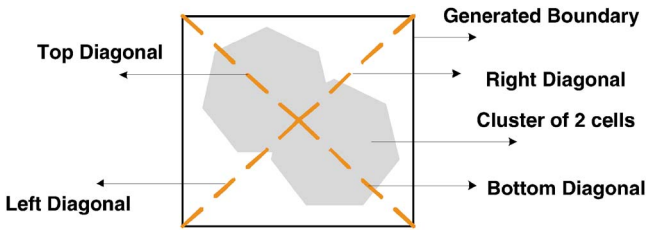


Figure 15. An example illustrating the left, right, top, and bottom diagonals on a cell cluster enclosed by a cell boundary.

Table I. Conditions for Cell Count within Clusters.

Test 1 and Test 2	Cell count
(BCT–BCZ) or (BRT–BRZ)=30–60 and CCFP=35–75	2
(BCT–BCZ) or (BRT–BRZ)=60–90 and CCFP=75–115	3
(BCT–BCZ) or (BRT–BRZ)=90–120 and CCFP=115–155	4
(BCT–BCZ)>120 or (BRT–BRZ) >120 and CCFP >155	>4

count within each cluster. As seen in Table I, a combined effect of conditions from Test 1 and Test 2 can assist in obtaining precise counts within clusters.

Thus, a combination of Test 1 and Test 2 assists in grouping clusters of two cells, three cells, four cells, and so on. The total cell count could then be computed by cumulatively adding the single cell count with the cluster count revealing the numbers of cells within each cluster, as shown in Eq. (2).

$$\text{Total} = (1s) + (2s \times 2) + (3s \times 3) + (4s \times 4) + \cdots . \qquad (2)$$

In addition to these statistics some efforts were directed toward another interesting feature—cell orientation. The cell cluster orientation was determined from the diagonal pixels versus the pixel values corresponding to the diagonal. The diagonals as previously mentioned were obtained using the boundary dimensions enclosing each cell cluster. Thus, a plot of diagonal pixels versus foreground pixel intensities could indicate cluster orientations; for example, if the cluster is oriented either toward the right or the left or is perfectly upright (either horizontal or vertical). As seen in the plot in Figure 16 top and bottom diagonals have a higher pixel count and fewer gaps in the plot than the right and left diagonals. This indicates that cluster 1 as shown in Figure 16 is oriented (tilted) toward the upper left side of the generated boundary. Similarly, as evident in the plot in Figure 17, cluster 2 is oriented toward the upper right side of the generated boundary since the top and bottom diagonals have more gap than the left and right ones. Figure 18 shows a plot with uniform gaps on all four diagonals indicating that cluster 3 is either oriented to be perfectly horizontal or perfectly vertical and in the exact center of the generated boundary. Thus, the above method could give us a clue about cluster

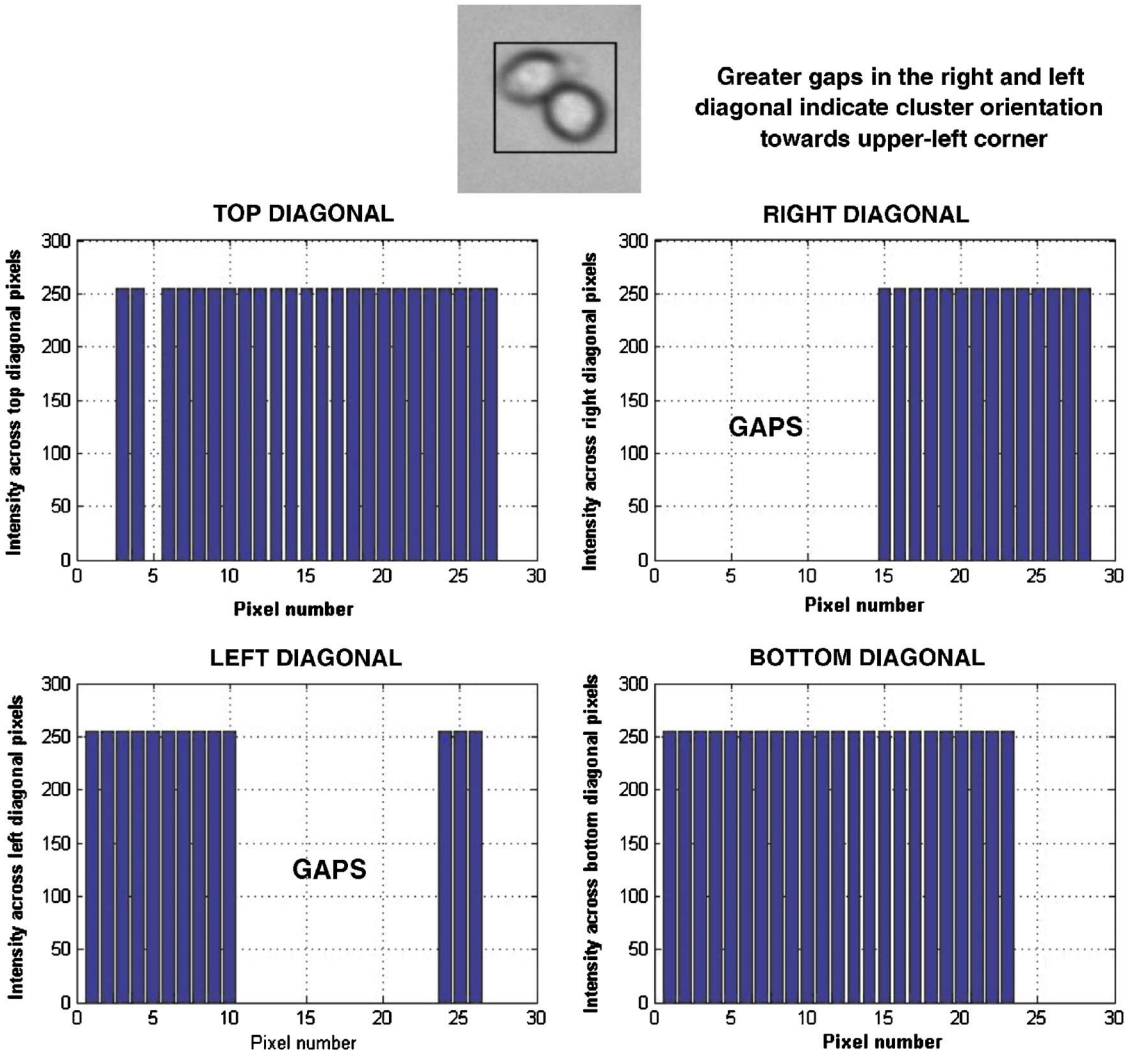


Figure 16. Plot of diagonal pixels vs foreground pixel intensities to indicate cluster orientations towards upper-left side of the generated cell boundary.

orientations as well. However, this is a preliminary test requiring more elaborate analysis to conclusively determine the cluster/cell orientations. No further efforts were directed toward this step.

(5) Final Quantification. The final step in the algorithm is to quantify and display the population statistics along with the processed image. The population statistics include single cell count, cell cluster count, cell count within each cluster, and the total cell count. The approximate cell size and cell position could also be determined from generated boundary dimensions and boundary positions, respectively.

PROCESSING RESULTS

The proposed algorithm is implemented in C/C++ and the performance is evaluated on a Pentium 4 3 GHz processor on a Windows environment. A Windows dialog based GUI as shown in Figure 19 was built using microsoft foundation classes (MFC) to achieve complete processing with a few button clicks. Most of the buttons as seen in Fig. 19 are selfexplanatory. The user is prompted with comments that appear in the edit window of the GUI after processing is complete.

The robustness of the designed algorithm is tested by processing a variety of cell images that vary in image sizes, cell sizes, intensity distributions, acquisition conditions, and noise levels. To this effect tests on the proposed algorithm were categorized as follows:

- *Case 1. Leukemia cells under bright field*

  Figures 20(a)–20(f) illustrate processed leukemia cell images acquired under bright field with white boundaries representing single cells and black boundaries representing cell clusters (refer to the image Acquisition subsection of this article for details on the test images). Test images shown in Figures 20(a) and 20(b) are acquired under ideal lighting conditions with defined cell edges. Unlike the ideal acquisition conditions that were seen in Figs. 20(a) and 20(b) the test images in Figs. 20(c) and 20(d) has nonuniform lighting conditions during acquisition, resulting in varying back-

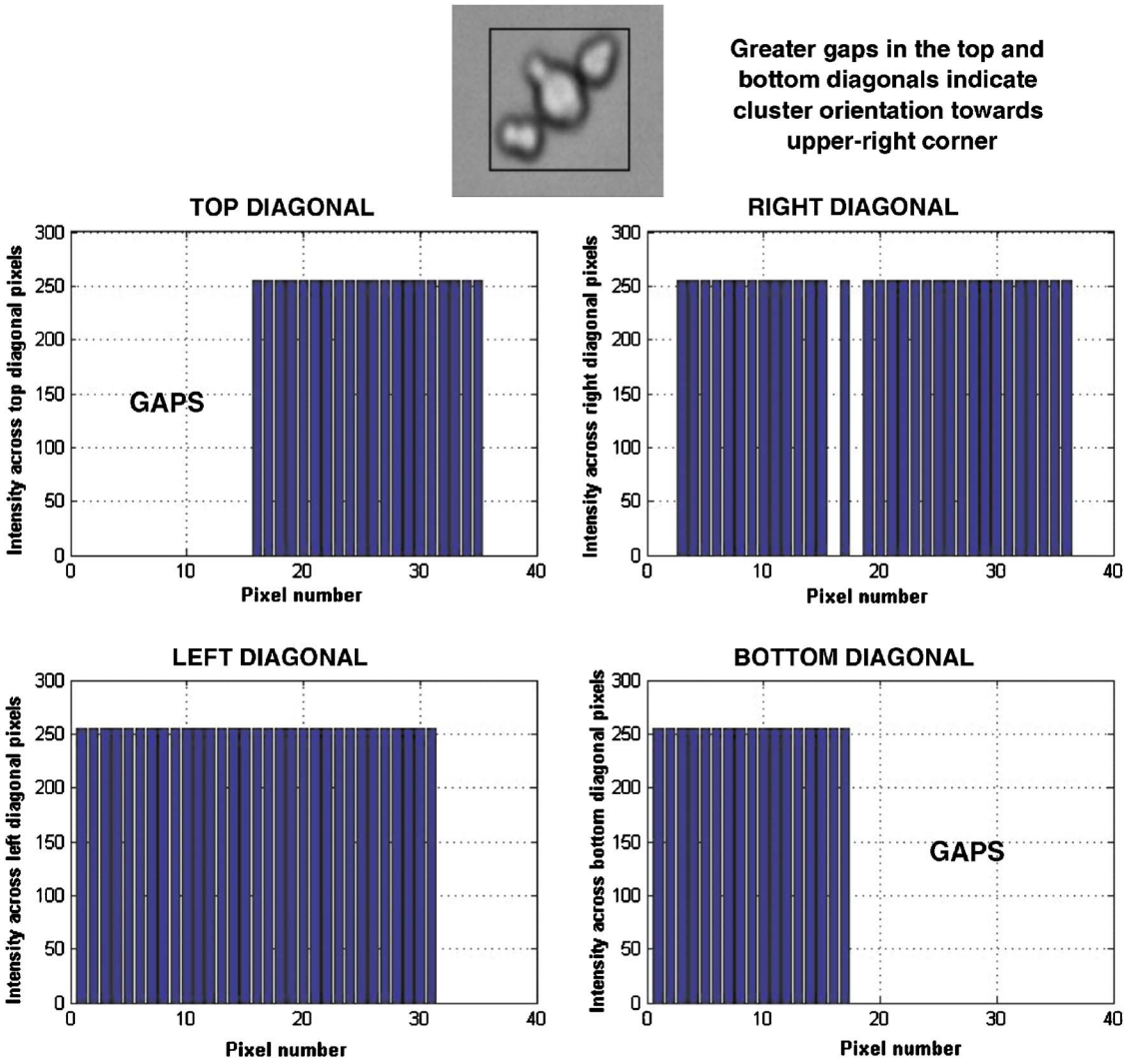


Figure 17. Plot of diagonal pixels vs foreground pixel intensities to indicate cluster orientations toward upper-right side of the generated cell boundary.

ground intensities, giving a gradient effect on the image. The processing results are successful, even on images with overlapping foreground and background pixel intensities as seen in Figs. 20(e) and 20(f). Effective elimination of background pixels under nonuniform illumination was obtained by recursively partitioning the image by strategically eliminating unwanted partitions and collecting the interested partitions in greater resolution. The image was initially horizontally partitioned into four main splits. Then, pixel intensities in each partition are summed and compared to an optimum adaptive threshold value for each subimage. The optimum threshold is computed by investigating its histogram since smaller image regions are more likely to have approximately uniform illumination. Thus, alternating horizontally and vertically, the division process repeats until most of the background is eliminated from the foreground. The looping value for recursive horizontal and vertical partitions needs to be strategically chosen to maintain reasonable accuracy, resolution, and processing speed. As mentioned earlier, details on the recursive segmentation process and the other processing steps in the cell identification stage were given in our earlier publications.[39,40] Table II provides detailed population statistics that include single cell count and cluster count including count of cells within each cluster in groups of two, three, four, or greater than four. The approximate cell area and the cell position could also be determined using the boundary box dimensions. However, the average cell area computed using the boundary box dimensions is only an approximate estimate of the cell size area since the boundary box is 2%–3% larger in extent than the true cell area.

- *Case 2. Leukemia cells under dark field*

The designed algorithm was also tested on images acquired in dark field. By using different stains, one can preferentially stain certain cell components, such as a nucleus or a cell wall, or the entire cell. The most basic reason that cells are stained is to enhance visualization of the cell or certain cellular components under a mi-

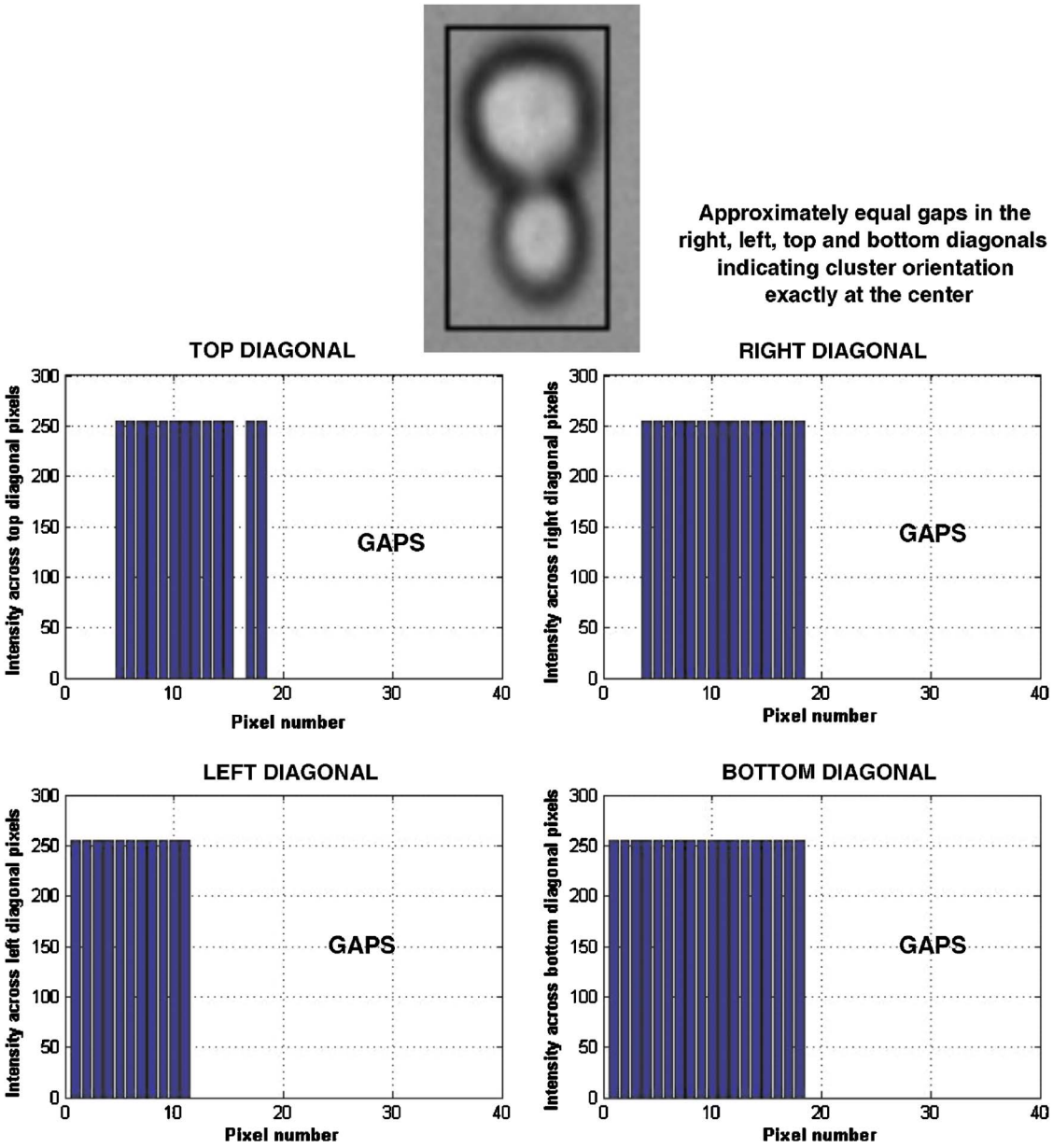


Figure 18. Plot of diagonal pixels vs foreground pixel intensities to indicate cluster orientations at the center of the generated cell boundary.

Table II. Population statistics of leukemia cells under bright field.

	Single cell count	Cell cluster count				Total count
		2 s	3 s	4 s	>4 s	
3e10	130	4	0	0	0	138
FMC44	105	4	2	1	0	123
3t39brev	0	0	1	0	1	7+
CD44i	1	0	0	1	0	5
MMB1	19	0	0	0	0	19
MMB3	29	3	0	0	0	35

croscope. Cells may also be stained to highlight metabolic processes or to differentiate between live and dead cells in a sample. Figures 21(a)–21(e) illustrate the processed leukemia cell images acquired under dark-field images with white boundaries representing single cells and grey boundaries representing cell clusters (refer to the Image Acquisition subsection of this article for details on the test images). The leukemia cells used for testing the designed algorithm were fluorescently stained using red or green Alexa dye. A stained cell in the image processing context appears as a bright colored spot that has no clean border with a relative intensity higher than its immediate neighbor. Accurate processing and classification of fluorescent spots is challenging since these images are typically low in contrast, high in

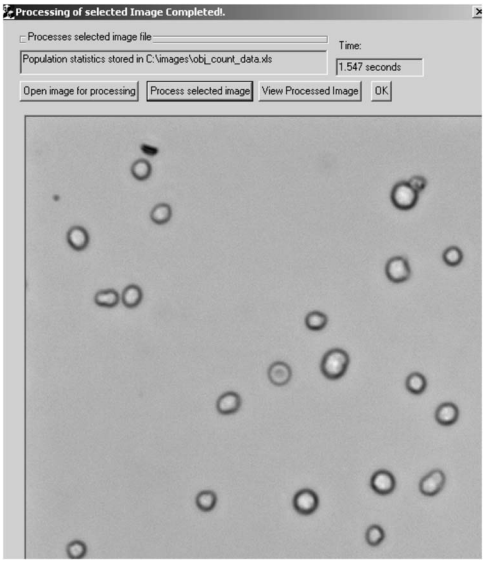


Figure 19. Graphical user interface of the proposed algorithm.

noise, and have no defined edge. Figure 21 illustrates processed images of stained images with scattered, jagged edges, and no defined shape resulting from ruptured cell walls from staining experiments. In addition, as observed in Figs. 21(a) and 21(b) there are impartially stained cells appearing as a colored spot with no defined size or edge. This could be due to either improper staining procedures in the experimental stage or to the absence of a certain proportion of cellular component being detected. The algorithm is designed such that it detects all the fluorescent spots in the field of view whether it is an entire stained cell or a partially stained cell. It is left to the discretion of the biologists to pick only completely stained cells or complete population statistics for further analysis. The designed algorithm processes the color channels (RGB) and luminance channels (Y) to ensure that any fluorescent spot missed in one channel is identified in another channel. Table III provides population statistics that includes isolated cell and cluster count.

The designed algorithm also extracts the relative staining intensity to inspect multicolored staining effect as in Fig. 21(e). The RGB intensity contribution of Fig. 21(e) is shown in Figure 22. The red component in Fig. 22 is five times greater than the green component and 20 times greater than the blue component, indicating greater affinity of the host molecule toward the red colored fluorescent marker. This feature opens up other possibilities to concurrently image multiple colored, spectrally overlapping marker proteins within living cells.[41,42] This ability to multiplex or detect multiple labels in the same experiment is both time- and cost-effective and improves accuracy.[42] Although approximate manual detection of relative intensity is possible, the task is subjective and does not always work well when the contrast of a cell image is poor. The results illustrated in Fig. 21 prove that the algorithm is sensitive

Table III. Population statistics of leukemia cells under dark field.

	Single cell count	Cell cluster count				Totalcount
		2 s	3 s	4 s	>4 s	
MMB3 (red)	15	5	0	0	0	25
SG28RFLt2 (green)	9	0	0	1	0	13
MMB1 (red)	10	0	0	0	0	10
CD44i (red)	1	1	0	0	0	3
3t39brev (red+green)	2	0	0	0	0	2

enough to filter noisy pixels and identify all the fluorescently stained cells, even the ones that are invisible to the naked eye under normal resolution.

- *Case 3. Leukemia cell lines under hemocytometer*

  The designed algorithm was also tested on leukemia cell lines on a conventional hemocytometer. The hemocytometer resembling a common microscopic slide has a chamber area and depth of known dimensions. The etched chamber resembles a grid that assists in extracting the cell count and computing the cell concentration or cell density in a specific volume of fluid. The hemocytometer is often used in a laboratory setup to quantify leukemia cell count and thereby determine the viability of leukemia cell samples for experiments.[43,44] Figures 23(a)–23(c) illustrate processing results on daudi cell lines on the hemocytometer grid (refer to the Image Acquisition subsection of this article for details on the test images). Processing and selectively extracting and classifying cells is effectively performed on diluted samples with fewer cells, as in Fig. 23(c) or dense samples with a large number of isolated cells and cell clumps, as in Fig. 23(a). Nonuniform background spread and dust particles on the hemocytometer slide, as seen in Fig. 23(a), further added to the processing challenge. Figure 23(b) illustrates an example of cells with varying appearance within the same sample. Ruptured cell walls have resulted in giving a hazy appearance to a large majority of cells with defined edges on nonruptured cells. Figure 23(b) is successful in obtaining accurate population statistics on images with varying cell appearance within the same sample.

The designed algorithm introduces an additional processing step (grid removal) after the median filtering step to process leukemia cells under the hemocytometer. This step is necessary because the grids are not the object of interest, but a higher intensity could mask the pixels corresponding to cells. Additionally, a savings of processing time could be achieved by only processing pixels of interest. The grids appear as sharp spikes equidistant from each other on the intensity profile that are removed and superimposed in the final stage of processing. The area of interest (AOI) for processing the hemocytometer cell images is also determined from the spike positions. Table IV provides population statistics that include single cell and cluster count.

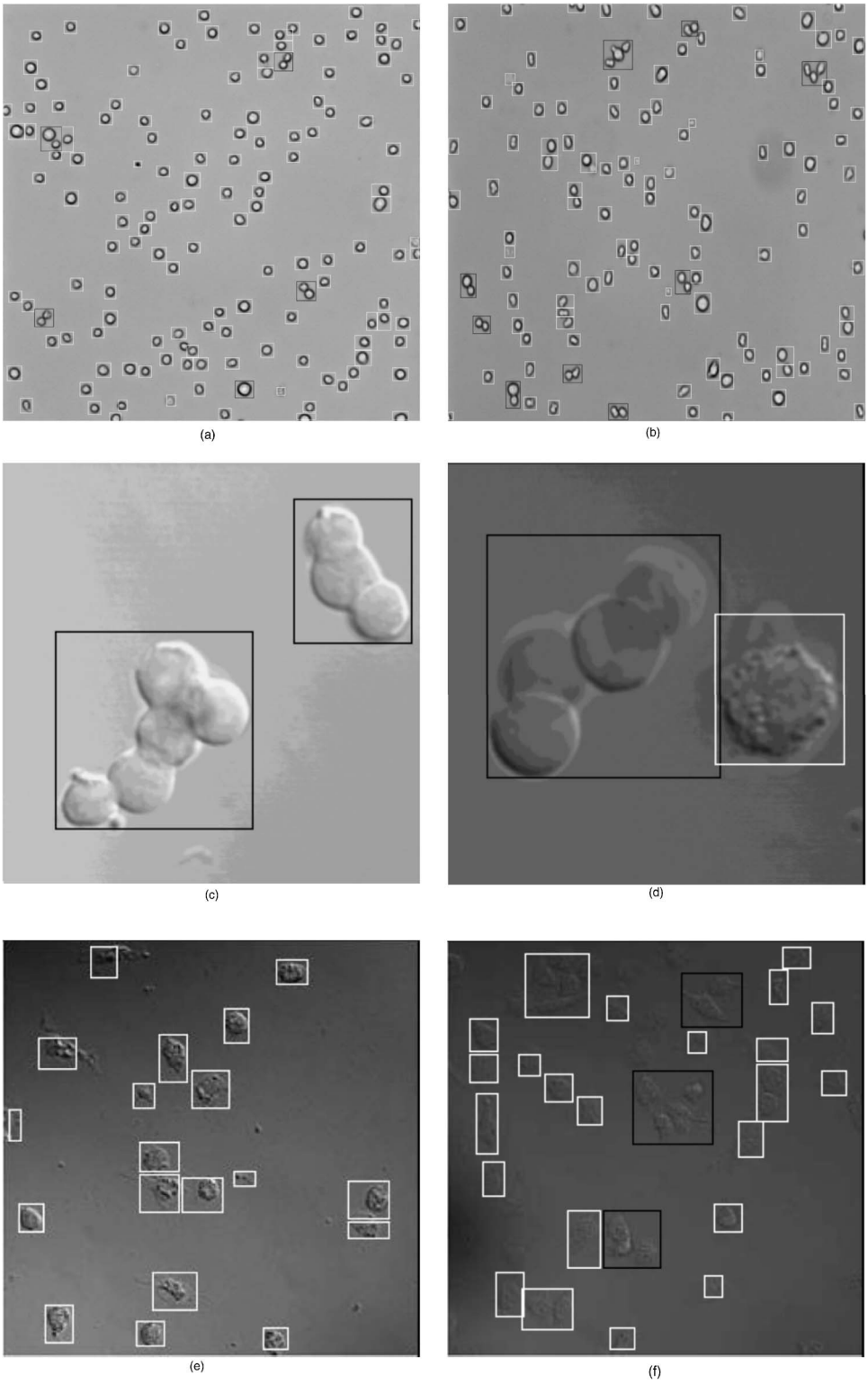


Figure 20. Processed images of leukemia cell types under bright field. White boundaries representing single cells and black boundaries representing cell clusters.

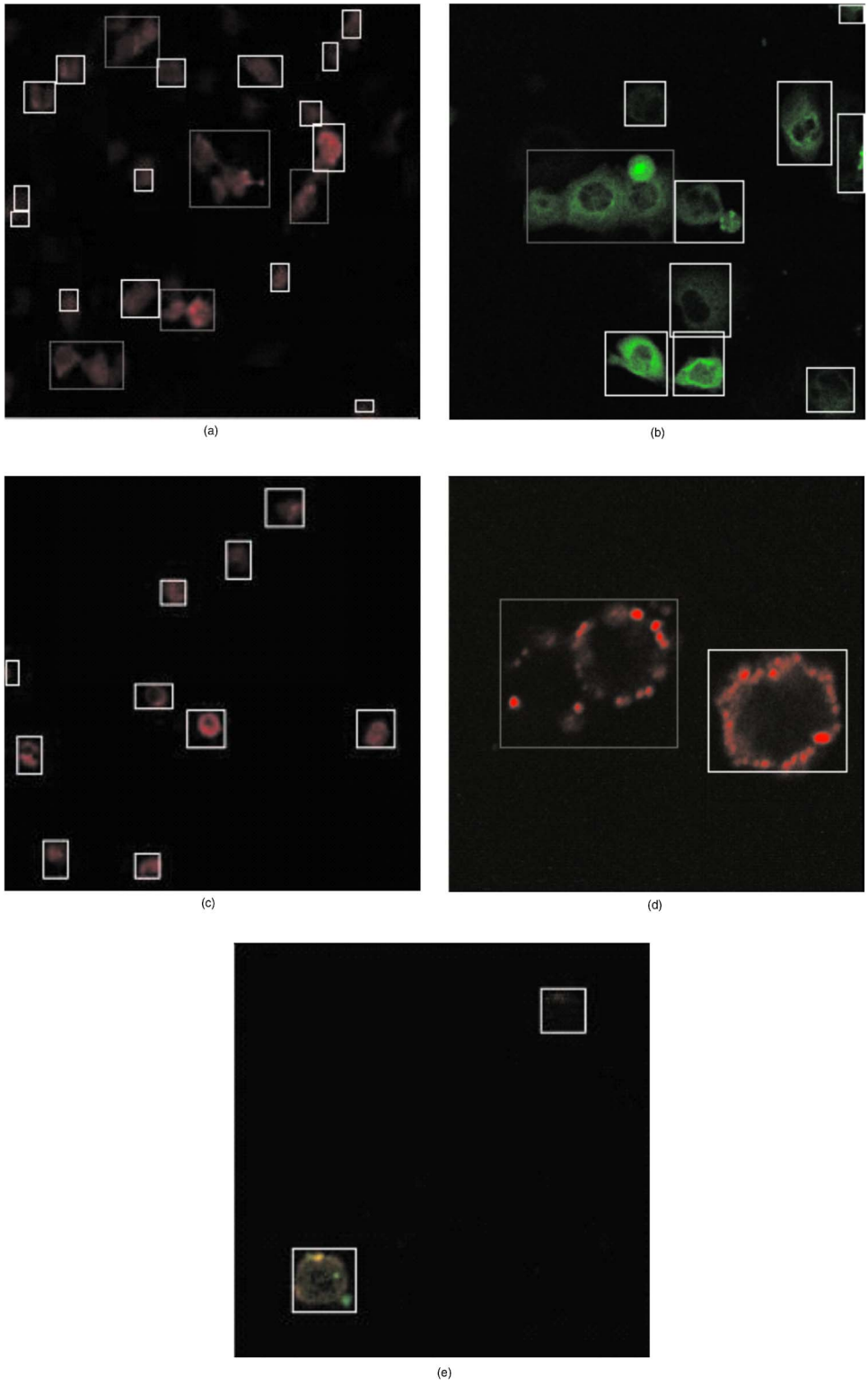


Figure 21. Processed images of leukemia cell types under dark field. White boundaries representing single cells and gray boundaries representing cell clusters.

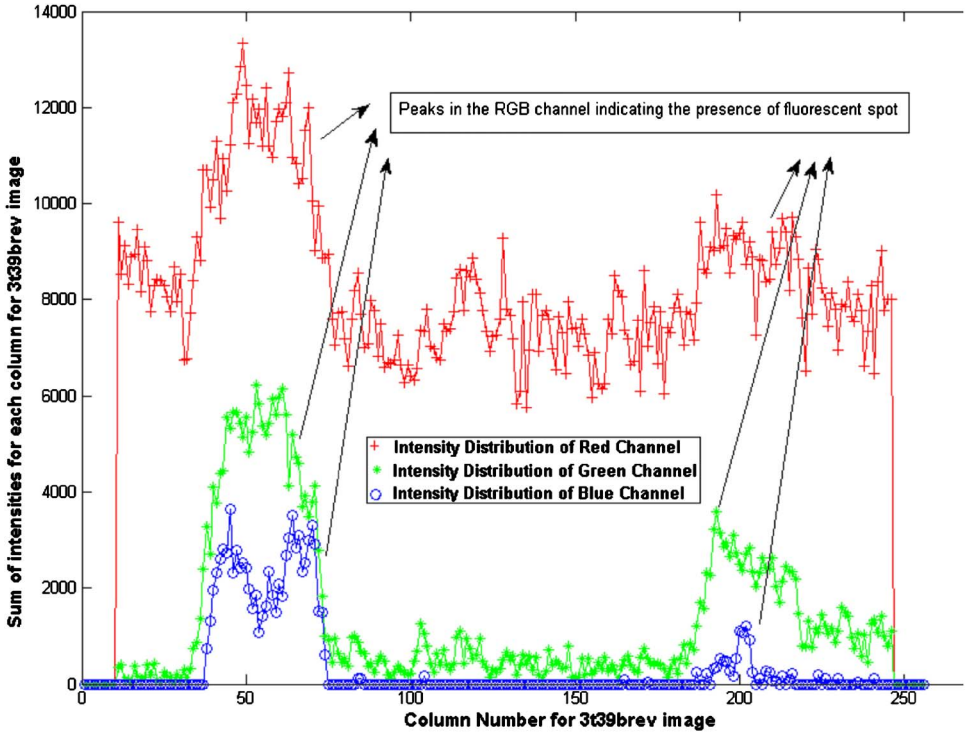


Figure 22. Relative intensity distribution plot of RGB channels of multicolored fluorescently stained 3t39brev cells.

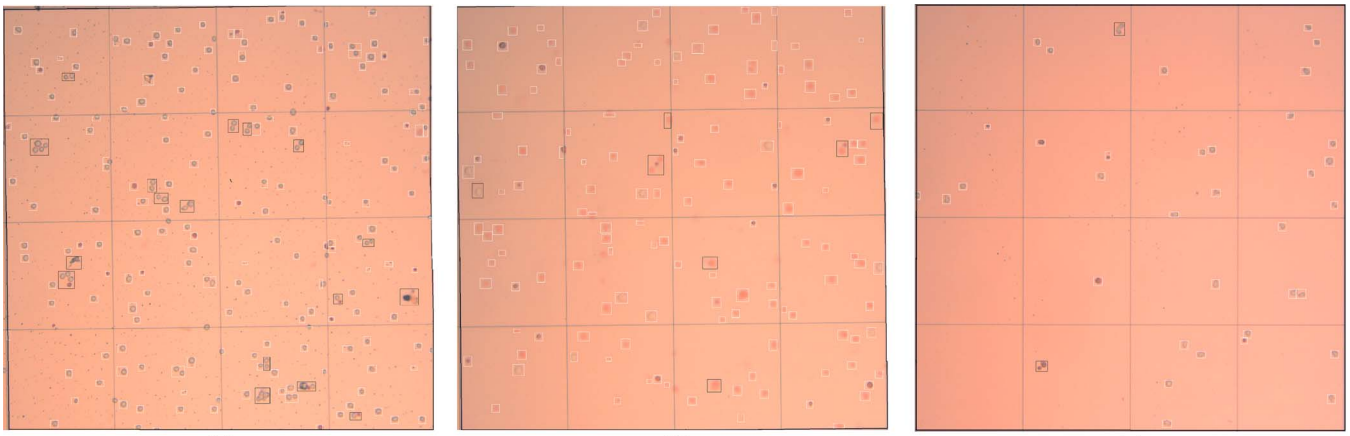


Figure 23. Test images of leukemia cell lines on a conventional hemocytometer. White boundaries representing single cells and black boundaries representing cell clusters.

PERFORMANCE ANALYSIS

It is difficult to carry out an objective comparison against different cell analysis algorithms because many algorithms are developed specific images and for specific tasks.[45] In other words, algorithms may not always work with types of cell images other than the ones they were designed for. For example, it would be hard to objectively compare the Nilsson and Heyden[27] algorithm developed for three-dimensional confocal images against the proposed algorithm because the Nilsson and Heyden algorithm makes the fundamental assumption that there are no overlapping cells since leukocyte counts are always performed in the monolayer part of the slide. However, to analyze the performance of the proposed techniques, comparisons were made against some recently published techniques and commercial image analysis tools using the manual techniques as obtained by trained biologists as a good standard.

Error Analysis

An error analysis is performed on the proposed algorithm by comparing the automated count and the manual count obtained by a biologist who arrived at these values after three repetitions. The average of the three repetitions was chosen

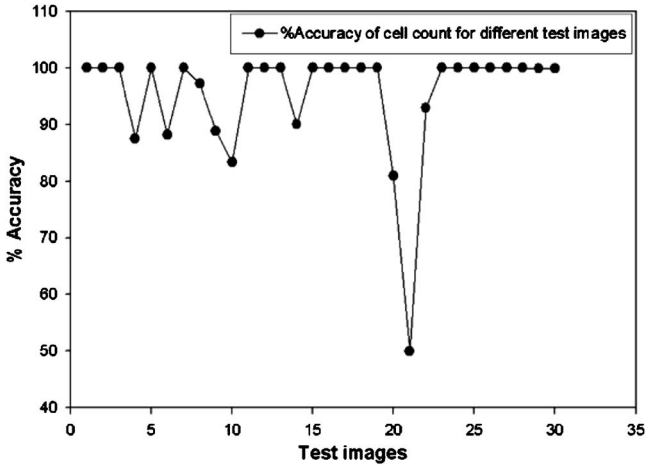


Figure 24. Percentage accuracy plot for all test images computed from the counting offsets.

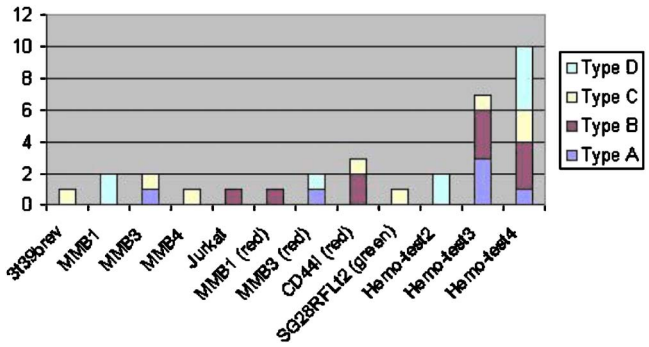


Figure 25. Classification of offsets in cell count into different error categories for different test images.

as a final manual count. In most cases, the count after three repetitions remained the same. Manual counts of all images were made by manual inspection of each image on the computer. Manual counts on hemocytometer test images were, however, performed by observing the cell sample on a microscope and using a ticking device. The manual and automated count for all test cases, along with the counting offset, is computed by taking the difference between the manual and automated count. It was observed that a maximum counting offset of ten cells is observed from the proposed algorithm. Figure 24 shows the percentage accuracy for all the test images computed from the counting offset with the *x*-axis in the plot indicating the test images. The plot shows that the accuracy ranges between 80%–100%, in most cases, with the average accuracy being approximately 95%. It can also be observed that the percentage accuracy computation is dependent on the counting offset/manual cell count ratio, such that the higher this ratio the lower is the percentage accuracy despite a low counting offset. The CD44i-red test image contained stained clusters with extremely weak and random fluorescent spots with scattered and broken edges causing errors in the signature profiles to identify the total cell count within the cluster, resulting in low counting accuracy (50%). It could also be concluded that the designed algorithm provides a better estimate in obtaining the total

Table IV. Population statistics of leukemia cells lines daudi on a hemocytometer.

	Single cell count	Cell cluster count				Totalcount
		2 s	3 s	4 s	>4 s	
Hemo 1	339	12	2	2	1	377+
Hemo 2	271	4	1	0	0	282
Hemo 3	35	2	0	0	0	39

count of cells in a dense sample with a large number of cells than a sparse sample with fewer cells (less than ten cells in a sample). This acts to the advantage of biologists since a high counting accuracy to quantify a dense sample with a large number of cells relieves the biologist from cumbersome manual counts than on samples with small cell numbers.

Assessing the accuracy only based on total count values could lead to inaccurate comparisons. For example, let us say that the manual total count obtained for an image is 20 such that there are 18 isolated cells and one cluster of two cells. Let us imagine the proposed algorithm accounts for population statistics with 17 isolated cells and one cluster of three cells. The total count from the proposed algorithm actually matches the manual count perfectly. However, the inaccuracy in identifying a cluster of three cells from a cluster of two cells is unaccounted if only the total count is compared. In order to avoid such imperfections all test images were compared, such that any offset in isolated cell or cluster count was further analyzed and categorized into four error categories:

- Type A If single isolated cells are classified as cell clusters or vice versa.
- Type B If a single cell or cell cluster is missed or ignored.
- Type C Total cell count within a cell cluster is incorrect.
- Type D Dust particles or other background artifacts are identified as cells and included in the count.

If there is no discrepancy in the isolated cell or cluster counts or the number of cells within the cluster, the processed image falls in the NO-ERROR category. No errors in identification and classification earn one point in the NO-ERROR category and zero in the other error categories. Counting errors on each test image is categorized accordingly, such that the number of points (earned) indicates the number of errors in that category. Figure 25 is used to illustrate the spread of counting errors in each category for all the test images (Case 1, Case 2, and Case 3). The errors in Type A category, denoting classification errors, are contributed from fluorescently stained images with jagged edges and irregular fluorescent stain that have caused ambiguity in classifying certain isolated stains as stained clusters or vice versa. Type B category errors, denoting identification errors, are contributed mostly from ignored cells exactly in the center of the hemocytometer grid or weak fluorescent spots that were ignored by the Y, R, G, and B channels. Type C category errors, denoting cluster analysis errors, are contributed

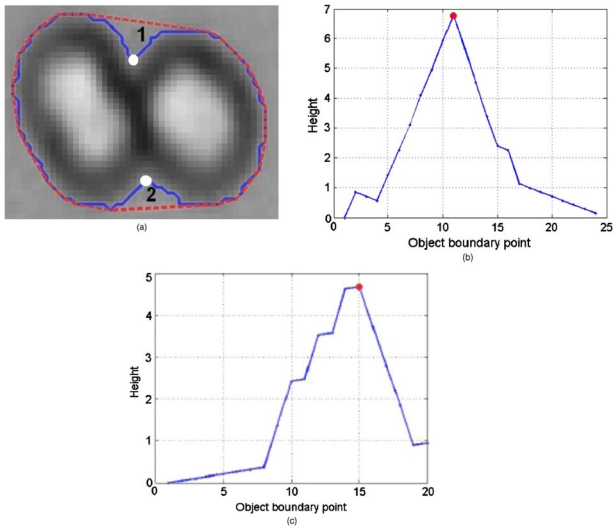


Figure 26. Convex hull with two concavities and their corresponding height profiles; two peaks in the height profile from the cell junctions in the two concavity regions indicates presence of two cells in a cluster.

from the uneven spread of fluorescently stained clusters, such that the 1D profile incorrectly classifies a cluster of three cells as a cluster of two cells. In addition, the proposed algorithm currently categorizes cell count within clusters in groups of two, three, four, and greater than four. Hence, a cluster with say five cells will fall in the greater than four category. The proposed algorithm gets a point in the Type C error category for all such cases. This is done because the exact cell counts for clusters greater than four cells are still unknown. Type D category errors, denoting sensitivity, are contributed from poor quality hemocytometer images that are unsuccessful in eliminating some noisy pixels thus adding to the count. These noisy pixels could arise from dust particles on the hemoslides that closely resemble the pixels of interest and could not be removed by the proposed algorithm. In spite of the errors generated, the proposed algorithm maintained an average counting accuracy of approximately 95% when manually compared with all test images. The counting offsets of seven to ten cells seemed negligible considering the image variety, image quality, and the nonideal acquisition conditions of the test images. However, comparisons with published methods further ascertain the robustness of the proposed technique.

Published Algorithms

The proposed algorithm is also analyzed and compared with some published algorithms to quantify population statistics of biological species. The total count and processing time comparisons with the proposed algorithm are discussed at the end of this section.

A concavity region is defined as any region that is bounded by a boundary arc and its corresponding convex hull chord.[30] These methods vary with respect to the technique for locating the concavity pixels and the cost function used to detect a split path. Two methods based on concavity analysis[30,31] (CA) are discussed here. In general, there are three sequential steps for concavity analysis: the detection of

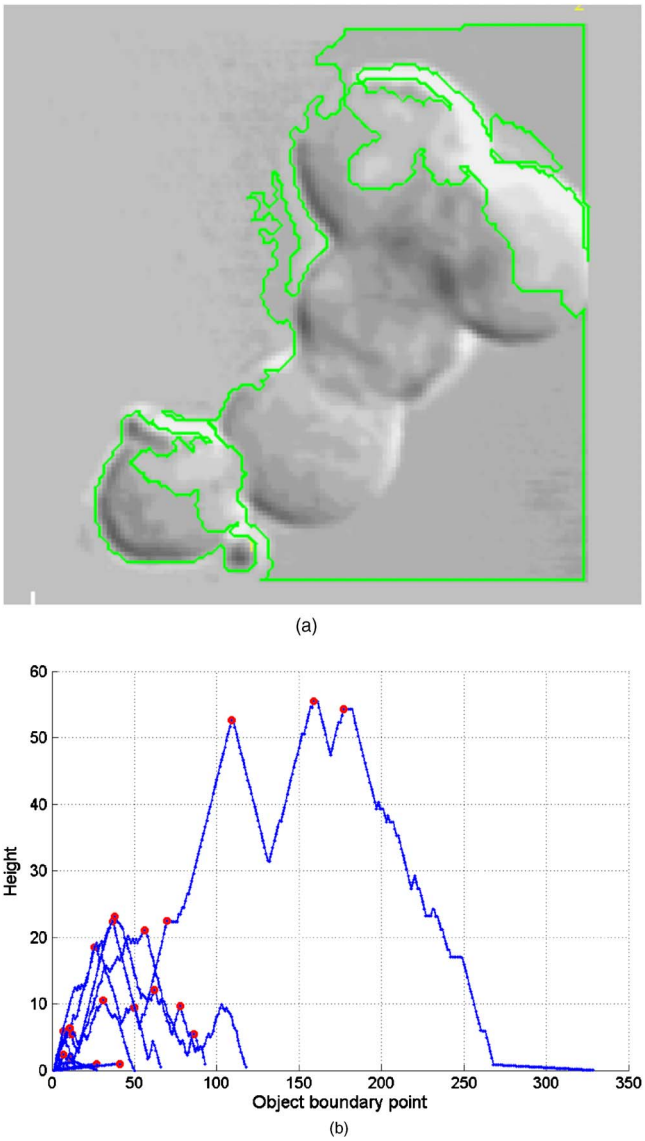


Figure 27. Cluster with five cells and their corresponding height profiles; too many peaks in the height profile making identification of cells within a cluster ambiguous.

concavity regions, the detection of candidate split lines, and the selection of best split line.[30] The best split line is recursively obtained until a specific stopping criterion is met.[30] The algorithm described by Choi[31] detects a concavity region by computing the percentage of the actual area of each traced object versus the area of the convex hull. A single cell is assumed with a convex hull very close to the boundary versus a cluster of cells with a convex hull showing appropriate concavities. Assuming each concavity as a cell junction a height profile is calculated using the perpendicular distances between the convex hull edge and actual cell edge points. After all concavities associated with the cell object are analyzed, detected peaks are counted to determine the number of cells within a cluster. Figures 26(b) and 26(c) indicate the height profile that shows two prominent peaks from the two cell junctions as seen in Fig. 26(a) indicating a cluster of two cells.[31] However, this proposed method could be ambiguous in identifying cell count within clusters with dif-

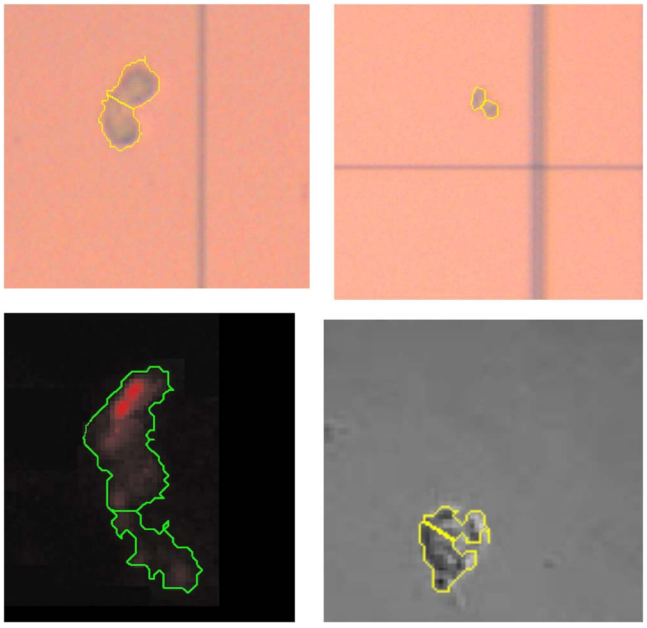


Figure 28. False clump splitting of isolated cells into two halves due to natural concavities in their boundaries.

fused edges, as seen in Figures 27(a) and 27(b) wherein the total number cells in the cluster does not correspond to the correct number of peaks in the height profiles. The Kumar et al.[30] algorithm computes concavity regions by searching for the pixel with the largest perpendicular distance from the corresponding convex hull, on each boundary arc to select a concavity pixel. A set of candidate split lines is obtained by joining all possible concavity pairs. The best split line from the set of candidate split lines is determined using a cost function. However, a false splitting rate due to boundary irregularities with large concavities, as seen in Figures 28(a) and 28(b), occur where a single cell with large concavities is split into two halves (assuming it is a cluster). Cell-Profiler (CP) is open-source software that was recently developed to automatically identify and measure a variety of biological objects in images.[33]

Figures 29(a) and 29(b) illustrate processed results on selected test images where the count falls short by approximately 15–20 cells when compared to the manual count. Figure 29 shows a count in excess of +12 cells due to the fact that the CP has deduced the isolated stained cell with dispersed intensity spots as a stained cluster dividing the as-

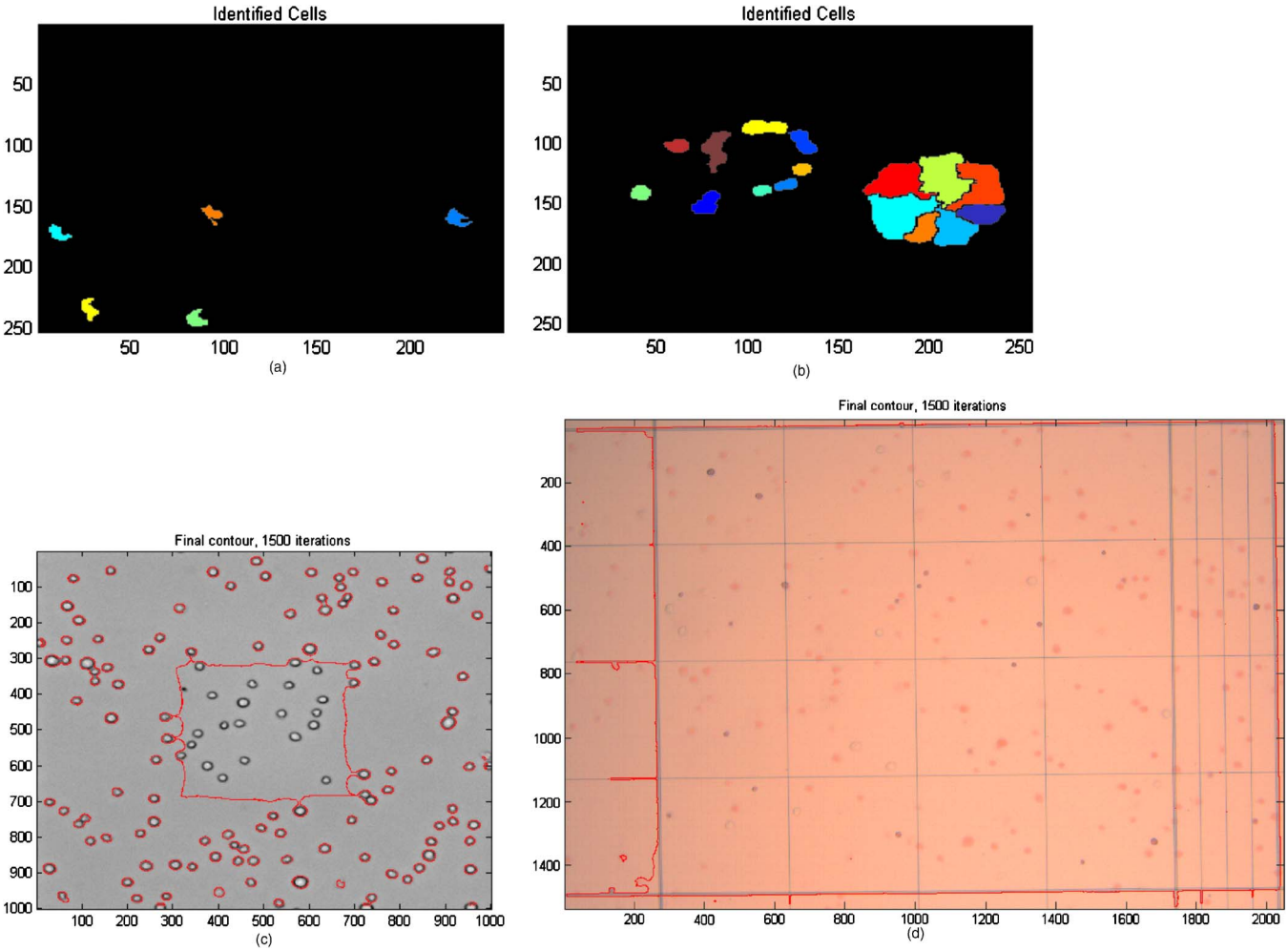


Figure 29. Processed images of selected test images using cell-profiler technique and variational active contours.

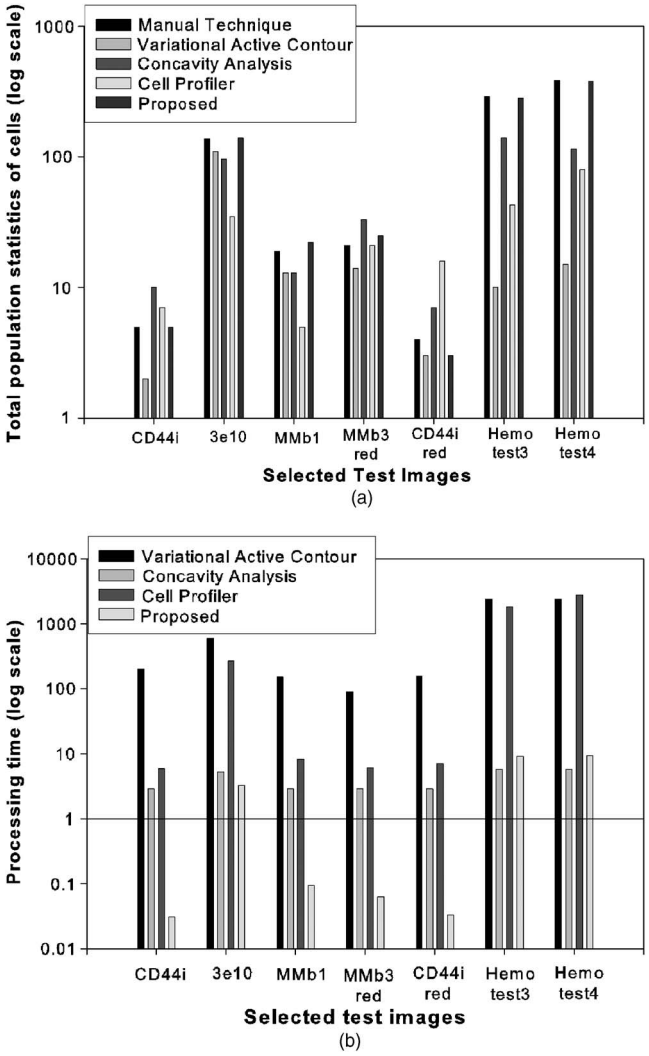


Figure 30. Population statistics and processing time (log scale) of the proposed technique compared against published algorithms (variational active contour, concavity analysis, cell-profiler) keeping manual technique as a gold standard.

sumed cluster into different sections. The proposed algorithm was also compared with a new variational level-set active contour formulation proposed by Li et al.[28] This new contour formulation consisted of an internal and an external energy term. The internal energy term penalizes the deviation of the level set function from a signed distance function, whereas the external energy term drives the motion of the zero level set to the desired image features such as object boundaries.[28] The processed results on variational active contour (VAC) formation demonstrate robustness of the contour formation to weak and low contrast boundaries. The closed contour from VAC formulation enclosing a cell is a better estimate of cell area than the proposed algorithm. However, the contour fails to move into clusters restricting the count accuracy to only isolated cells. Moreover, as seen in Fig. 29(c), the contour formation on FMC44 test image is abruptly terminated since the contour formation took ~24 min to process the FMC44 test image with 1500 iterations. The processed results from hemocytometer test images, as shown in Fig. 29(d), fail to form closed contours across cells even after 1500 iterations converging to forming contours around the grids.

Figure 30(a) compares the total count of cells (log scale) from published algorithms or techniques (VAC, CA, CP) with the proposed algorithm keeping manual count as the gold standard. The *y*-axis of the plot is drawn on log scale to improve visibility. The comparison of the population statistics was done on selected test images that vary in image quality and acquisition conditions. It was observed that a maximum counting offset of seven to ten cells from the proposed technique was observed in comparison to the manual technique versus a maximum counting offset of 62, 26, and 29 cells from VAC, CA, and CP, respectively. As previously discussed, the count statistics indicate minimum error with well-defined uniformly lighted and noise-free test images. However, the accuracy is severely affected with nonideal test images leading to false clump splitting or ambiguous height profiles as in CA. It was also observed that the published techniques indicated maximum offsets in count with hemocytometer test images (hemo-test3, hemo-test4) and fluorescent test images (*MMB3-red, CD44i-red*). These images are most commonly used in laboratories making the proposed algorithm an ideal choice for quantifying population statistics on such images. The processing time is also indicated in Fig. 30(b) (log scale) such that the proposed algorithm takes a maximum processing time of 9.4 s to process a hemo-test4 as against 24 s, 5.8 s, 27 s from VAC, CA CP, techniques, respectively. It can be observed that the hemocytometer test images from the proposed algorithm take approximately 9 s to process hemo-test4 image compared to concavity analysis that takes only ~5 s. However, the accuracy of count is less than 65% with concavity techniques for hemo test images than 90% with the proposed algorithm. Thus, the proposed algorithm demonstrates good potential in processing both ideal and nonideal images with an average accuracy of 91% and average processing time of 3 s.

Commercial Technique

The proposed algorithm is also compared with commercial image processing software with an automatic cell counting setup.[35] Several attempts were made on the commercial image analysis tool to alter thresholds and filter and smoothen the image to achieve the best results. Identified and classified cells are represented by contours enclosing cell boundary that are eventually quantified to obtain the final count. The commercial image analysis tool was comparable to the designed algorithm with only selected images. However, the efficiency drastically reduced when the image background was uneven or image quality was poor. Figures 31(a)–31(c) show the results of the commercial software on selected test images. As seen in Fig. 31(a), the processed results from the commercial software on leukemia cell type 3e10 provides a total count comparable to the proposed and manual technique. However, the processed results on CD44i cell type and the daudi cell lines on the hemocytometer are almost completely masked in the presence of grids. In addition, as

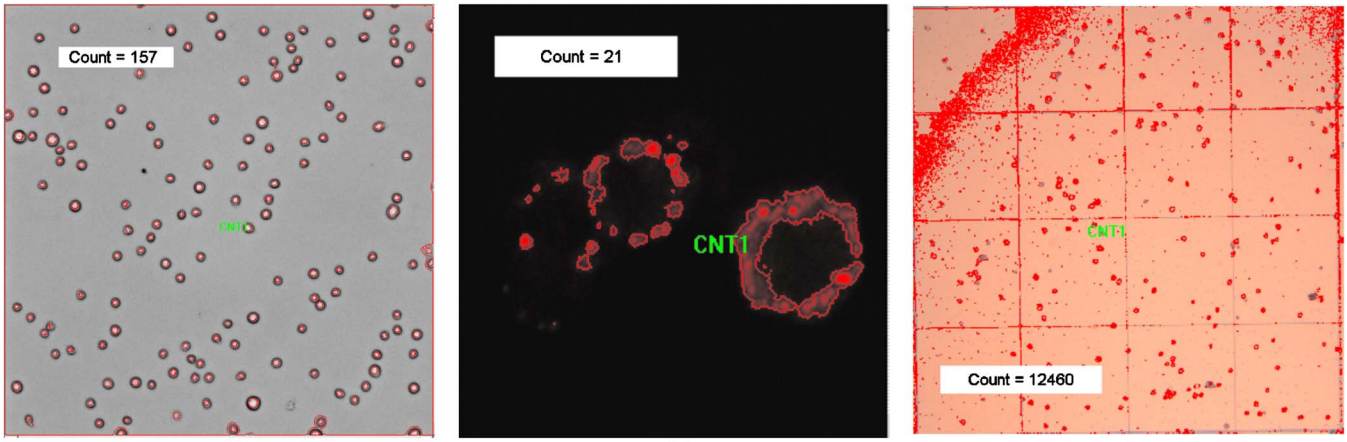


Figure 31. Processed image of test images using commercial software[35] with automatic cell counting setup.

seen in Fig. 31(c), the background noise leads to false positives, which results in an incorrect count.

CONCLUSION

An efficient and high-throughput leukemia cell count and cluster classification algorithm has been presented. Signature profiles are used to evaluate and determine cell count within each cell cluster. The validity of the algorithm was tested with under bright field, leukemia cells under dark field, and on a conventional hemocytometer. The performance of the algorithm was also accessed with poor contrast images and noisy images typically arising from poor lighting conditions during acquisition or artifacts from experimental procedures. Comparisons with the concavity analysis, CellProfiler, variational active contours, and commercial image analysis tool further exhibits the robustness of the designed algorithm to a variety of images. The proposed algorithm demonstrates good potential in processing both ideal and nonideal images with an average accuracy of 95% and average processing time of 3 s. The designed algorithm proves to be efficient, accurate, and reliable that could improve overall throughput of experiments.

ACKNOWLEDGMENTS

The authors would like to thank the reviewers for their valuable comments and suggestions in improving this article and NSERC Strategic Project (2002–2007) and the Department of Electrical and Computer Engineering at University of Calgary for their financial support. The authors would also like to thank Linda Pilarski, Department of Oncology, and the University of Alberta for providing the leukemia test image samples and Colin Dalton and Daniella Spiller from the University of Calgary for providing hemocytometer leukemia cell images and assisting in manual counting procedures.

REFERENCES

[1] National Cancer Institute, *What you need to know about leukemia*, Available: http://www.cancer.gov/cancerinformation/doc.aspx?viewid=57b3abc6- 4b52- 41b0- 8762- 9372062313de
[2] L. M. Franks and N. M. Teich, in *Introduction to the Cellular and Molecular Biology of Cancer*, 3rd ed. (Oxford University Press, Oxford, UK, 1997) pp. 251–254.
[3] F. Athanassiadou, M. Kourti, T. Papageorgiou, M. Stamou, A. Makedou, and A. Boufidou, "Severe hyperlipidemia in a child with acute lymphoblastic leukemia treated with l-asparaginase and prednisone," Pediatrics International **46**, 743–744 (2004).
[4] N. Abramson and B. Melton, "Leukocytosis: Basics of clinical assessment", Am. Fam. Physician **62**, 2053-2060 (2000).
[5] B. Nilsson and A. Heyden, "Segmentation of dense leukocyte clusters", *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA01)*, (IEEE, Piscataway, NJ, 2001), pp. 221–227.
[6] D. K. Yip and N. Auersperg, "Structural and functional heterogeneity of the surface of rat leukemia cells", J. Cell Biol. **63**, 109–124 (1974).
[7] I. Francis, M. Adeyanju, S. George, T. Junaid, and U. Luthr, "Manual versus image analysis estimation of pcna in breast carcinoma", Anal. Quant. Cytol. Histol. **22**, 11–16 (2000).
[8] C. Macaulay, "A comparison of some quick and simple threshold selection methods for stained cells", Anal. Quant. Cytol. Histol. **10**, 134 (1988).
[9] R. Umesh, "A technique for cluster formation", Pattern Recogn. **21**, 393 (1988).
[10] J. F. Hadon, "Generalized threshold selection for edge detection", Pattern Recogn. **21**, 195–203 (1988).
[11] R. Ohlander, "Picture segmentation using a recursive region splitting method", Comput. Graph. Image Process. **8**, 313 (1978).
[12] C. Garbay, "An iterative region-growing process for cell image segmentation based on local color similarity and global shape criteria", Anal. Quant. Cytol. Histol. **8**, 25 (1986).
[13] G. Haussmann, "A region extraction approach to blood smear segmentation", Comput. Vis. Graph. Image Process. **25**, 133 (1984).
[14] H. Aus, A. Ruter, V. T. Meulen, U. Gunzer, and R. Nurnberger, "Bone marrow cell scene segmentation by computer-aided color cytophotometry", J. Histochem. Cytochem. **25**, 662–667 (1977).
[15] J. Brenner, T. Necheles, I. Bonacossa, R. Fristensky, B. Weintraub, and P. Neurath, "Scene segmentation techniques for the analysis of routine bone marrow smears from acute lymphoblastic leukemia patients", J. Histochem. Cytochem. **25**, 601–613 (1977).
[16] H. Harms, H. Aus, M. Haucke, and U. Gunzer, "Segmentation of stained blood cell images measured at high scanning density with high magnification and high numerical aperture optics", Cytometry **7**, 522–531 (1986).
[17] G. Landeweerd, E. Gelsema, J. Brenner, W. Selles, and D. Zahniser, "Pattern recognition of nucleated cells from the peripheral blood", Pattern Recogn. **16**, 131–140 (1983).
[18] D. Wermser, G. Haussmann, and C. Liedtke, "Segmentation of blood smears by hierarchical thresholding", Comput. Vis. Graph. Image Process. **25**, 151–168 (1984).
[19] S. Poon, R. Ward, and B. Palcic, "Automated image detection and segmentation in blood smears", Cytometry **13**, 766–774 (1992).
[20] S. J. Lockett and B. Herman, "Automatic detection of clustered fluorescently-stained nuclei by digital image-based cytometry", Cytometry **17**, 1–12 (1994).
[21] N. Malpica, C. O. de Solorzano, J. J. Vaquero, A. Santos, I. Vallcorba, J. M. García-Sagredo, and F. del Pozo, "Applying watershed algorithms to

the segmentation of clustered nuclei", Cytometry **28**, 289–297 (1997).
[22] P. Sjöstrom, B. Frydel, and L. Wahlberg, "Artificial neural network-aided image analysis system for cell counting", Cytometry **36**, 18–26 (1999).
[23] G. D. C. G. Loukas and B. Vojnovic, "Automated segmentation of cancer cell nuclei in complex tissue sections", Proc. SPIE **4158**, 188–198 (2001).
[24] Y. Ma, R. Dai, and L. Li, "A counting and segmentation method of blood cell image with logical and morphological feature of cell", Chin. J. Electron. **11**, 53–55 (2002).
[25] W. Hsu, M. Lee, and B. Fang, "On the accurate counting of tumor cells", IEEE Trans. Nanobiosci. **2**, 94–103 (2003).
[26] B. Swolin, P. Simonsson, S. Backman, I. Lofqvist, I. Bredin, and M. Johnsson, "Differential counting of blood leukocytes using automated microscopy and a decision support system based on artificial neural networks evaluation of diffmaster octavia", Clin. Lab Haematol. **25**, 139–147 (2003).
[27] B. Nilsson and A. Heyden, "Segmentation of complex cell clusters in microscopic images: Application to bone marrow samples", Cytometry **66A**, 24–31 (2005).
[28] C. Li, C. Xu, and C. G. M. Fox, "Level set evolution without re-initialization: A new variational formulation", *Proc. 2005 IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR05)*, (IEEE, Piscataway, NJ, 2005).
[29] B. Prasad, S. Du, W. Badawy, and K. Kaler, "A real-time multiple-cell tracking platform for dielectrophoresis (dep)-based cellular analysis", Meas. Sci. Technol. **16**, 909–924 (2005).
[30] S. Kumar, S. Ong, S. Ranganath, T. Ong, and F. Chew, "A rule-based approach for robust clump splitting", Pattern Recogn. **39**, 1088–1089 (2005).
[31] J. Choi, "A high throughput cytometry algorithm for leukemia cells", M.Sc. Thesis (University of Calgary, Calgary, Canada, 2007).
[32] H. Wu, V. Hanson, and J. Gil, "Estimation of cell count from cell culture images", J. Imaging Sci. Technol. **50**, 187–192 (2006).
[33] M. Lamprecht, D. Sabatini, and A. Carpenter, "Cell profiler: free, versatile software for automated biological image analysis", BioTechniques **42**, 71–75 (2007).
[34] H. Russel, "Image analysis software makes the catch," Advanced Imaging Magazine (Sept. 2005); Available at http://www.advancedimagingpro.com/publication/article.jsp?pubId=1&id=1744.
[35] Image analysis software and products from Media Cybernetics; Available: http://www.mediacy.com/index.aspx?page=ImageProFamily(2005).
[36] Mcid elite; Available: http://www.imagingresearch.com/products/MCID.asp(2005).
[37] Dp25 digital camera, Olympus Camera Inc.; Available: http://www.olympusamerica.com/seg section/seg_digcam.asp (last accessed May 2008)
[38] R. C. Gonzalez and R. Woods, *Digital Image Processing*, 2nd ed. (Prentice Hall, Englewood Cliffs, NJ, 2002) p. 5555.
[39] J. Choi, B. Prasad, and W. Badawy, "Efficient cell counting and cluster classification algorithm for biological experiments", *Proc. 2006 IEEE Canadian Conference on Electrical and Computer Engineering - CCECE 2006* (IEEE, Piscataway, NJ, May 2006).
[40] B. Prasad, J. Choi, and W. Badawy, "An efficient cell analysis algorithm for fluorescently stained cell images", *Proc. 29th Canadian Medical Biological Engineering Conference* (CMBEC29), Vancouver, Canada, July 1–3, 2006.
[41] J. Davidsson, K. Paulsson, and B. Johansson, "Multicolor fluorescence in situ hybridization characterization of cytogenetically polyclonal hematologic malignancies", Cancer Genet. Cytogenet **163**, 180–183 (2005).
[42] L. Zhao, K. Hayes, and A. Glassman, "Enhanced detection of chromosomal abnormalities with the use of rxfish multicolor banding technique", Cancer Genet. Cytogenet **118**, 108–111 (2000).
[43] J. Berkson, T. B. Magath, and M. Hurn, "The error of estimate of the blood cell count as made with the hemocytometer", Am. J. Physiol. **128**, 309–323 (1939).
[44] G. Graldi, G. Metteucci, and R. Reverberi, "Leukocyte counting with the nageotte hemocytometer: Use of fluorescence microscopy", Transfusion Medicine and Hemotherapy **26**, 222–225 (1999).
[45] K. Rodenacker and E. Bengtsson, "A feature set for cytometry on digitized microscopic images", Anal Cell Pathol. **25**, 1–36 (2003).