[14] V. Sholukha, P. Salvia, I. Hilal, V. Feipel, M. Rooze, and S. Van Sint Jan, "Calibration and validation of 6 DOF's instrumented spatial linkage for biomechanical applications. A practical approach," *Med. Eng. Phys.*, vol. 26, pp. 251–260, 2004.
[15] A. Leardini, J. J. O'Connor, F. Catani, and S. Giannini, "Kinematics of the human ankle complex in the passive flexion; a single degree of freedom system," *J. Biomech.*, vol. 32, pp. 111–118, 1999.
[16] H. Woltring, R. Huiskes, A. de Lange, and F. Veldpaus, "Finite centroid and helical axis estimation from noisy landmark measurements in the study of human motion kinematics," *J. Biomech.*, vol. 18, pp. 379–389, 1985.
[17] A. Lundberg, I. Goldie, B. Kalin, and G. Selvik, "Kinematics of the ankle/foot complex: Plantarflexion and dorsiflexion," *Foot Ankle*, vol. 9, pp. 194–200, 1989.
[18] A. Lundberg, O. Svensson, C. Byland, I. Goldie, and G. Selvik, "Kinematics of the ankle/foot complex—Part 2: Pronation and supination," *Foot Ankle*, vol. 9, pp. 248–253, 1989.
[19] A. Cappozzo, F. Catani, A. Leardini, M. Benedetti, and U. Della Croce, "Position and orientation of bones during movement: Experimental artifacts," *Clin. Biomech.*, vol. 11, pp. 90–100, 1996.
[20] C. Reinschmidt, A. Van den Bogert, A. Lundberg, B. Nigg, N. Murphy, A. Stacoff, and A. Stano, "Tibiofemoral and tibiocalcaneal motion during walking: external vs. skeletal markers," *Gait Posture*, vol. 6, pp. 98–109, 1997.
[21] L. Lucchetti, A. Cappozzo, A. Cappello, and U. Della Croce, "Skin movement artifact assessment and compensation in the estimation of knee-joint kinematics," *J. Biomech.*, vol. 31, pp. 977–984, 1998.
[22] J. Cox and T. Hewes, "Normal talar tilt angle," *Clin. Ortho. Rel. Res.*, vol. 140, pp. 37–41, 1979.
[23] A. Aroua, B. Burnand, I. Decka, J. Vader, and J. Valley, "Nation-wide survey on radiation doses in diagnostics and interventional radiology in Switzerland in 1998," *Health Phys.—Radiat. Safety J.*, vol. 83, pp. 46–55, 2002.

Automated Segmentation, Classification, and Tracking of Cancer Cell Nuclei in Time-Lapse Microscopy

Xiaowei Chen, Xiaobo Zhou, and Stephen T. C. Wong*

***Abstract*—Quantitative measurement of cell cycle progression in individual cells over time is important in understanding drug treatment effects on cancer cells. Recent advances in time-lapse fluorescence microscopy imaging have provided an important tool to study the cell cycle process under different conditions of perturbation. However, existing computational imaging methods are rather limited in analyzing and tracking such time-lapse datasets, and manual analysis is unreasonably time-consuming and subject to observer variances. This paper presents an automated system that integrates a series of advanced analysis methods to fill this gap. The cellular image analysis methods can be used to segment, classify, and track individual cells in a living cell population over a few days. Experimental results show that the proposed method is efficient and effective in cell tracking and phase identification.**

***Index Terms*—Image analysis, phase identification, time-lapse fluorescence microscopy, tracking.**

Manuscript received September 20, 2004; revised September 1, 2005. The work of ST. C. Wong was supported in part by the HCNR under a Center for Bioinformatics Program Grant. *Asterisk indicates corresponding author.*

X. Chen and X. Zhou are with the HCNR Center for Bioinformatics, Harvard Medical School, Boston, MA 02115 USA and also with the Department of Radiology, Brigham & Women's Hospital, Boston, MA 02215 USA.

*S. T. C. Wong are with the HCNR Center for Bioinformatics, Harvard Medical School, Boston, MA 02115 USA (e-mail: stephen_wong@hms.harvard.edu) and also with the Department of Radiology, Brigham & Women's Hospital, Boston, MA 02215 USA.

Digital Object Identifier 10.1109/TBME.2006.870201

I. Introduction

Measuring cell cycle progress under different drug treatment conditions would improve our understanding of the biological mechanisms of oncological diseases and enhance the effectiveness of drug discovery and development. Recent advances in time-lapse fluorescence microscopy imaging have provided an important tool to study the cell cycle process under different conditions of perturbation. Existing computational methods, however, are rather limited in dealing with such time-lapse datasets, whereas manual analysis is unreasonably time-consuming and subject to observer variance. In this paper, a new method is presented to measure progress of the cell cycle. Our goal is to develop a computational system that can be used to identify phase changes of individual cells over time. This requires the ability to segment, classify, and track nuclei in cellular image sequences.

Accurate segmentation of cell nuclei is crucial in the measurement and tracking of cell cycle phases. For this application, we first use Ostu's method [10] to segment nuclei from the background and then deploy a watershed technique [12] to further separate touching nuclei. A nuclei-fragment merging method based on the shapes and sizes of nuclei is used to deal with the problem of over-segmentation [1], [8], [11]. In addition, pattern recognition techniques can be used to identify cell phases based on features extracted from cell nuclei. In our preliminary study [3], we found a K-nearest neighbor (KNN) classifier [6] with a set of seven features generated good results. We thus continue to explore the KNN classifier for cell phase identification.

Nucleus tracking in time-lapse microscopy data is a complex endeavor. Fig. 1 shows an example of the division of nuclei and changes of nucleus appearance, such as intensity and morphology, during mitosis. Active contour techniques [7] have been explored in the past for cell tracking in video-microscopy [9]. Active contour tracking, however, requires cells to be at least partially overlapped in different frames for successful tracking. This requirement cannot be met during cell division as the daughter nuclei are split into different directions and do not overlap with the parent nucleus in the previous image frame. Cells also undergo dramatic changes in phenotypic appearance during mitosis. In such cases, a correlation tracker [4] will fail to work also. Thus, we propose a new tracking technique to deal with these challenges derived from the time-lapse cellular image datasets.

II. Method

A. Nuclei Segmentation and Fragments Merging

In time-lapse fluorescence microscopy images, nuclei are bright objects protruding out from a relatively uniform dark background. They can readily be segmented by applying global thresholding [2]. In our system, the Otsu's algorithm [10] was used to perform image thresholding. The algorithm correctly segments most of the isolated nuclei, but fails to resolve the ambiguity caused by nuclei that overlapping or touching each other. To resolve the issue of touching nuclei, an improved watershed algorithm described in [12] was used. Watershed techniques, however, would cause over-segmentation of certain nuclei [11]. Subsequently, a postprocessing step is needed to correct and merge those over-segmented nuclei fragments.

In two–dimensional microscopy images, nuclei are usually elliptic objects with various degrees of ellipticity. The compactness is often used to describe the shape of the nuclei. Compactness is defined as the ratio of the square of the perimeter of the nucleus to the area of the nucleus

$$C = \frac{P^2}{4\pi \times A} \tag{1}$$

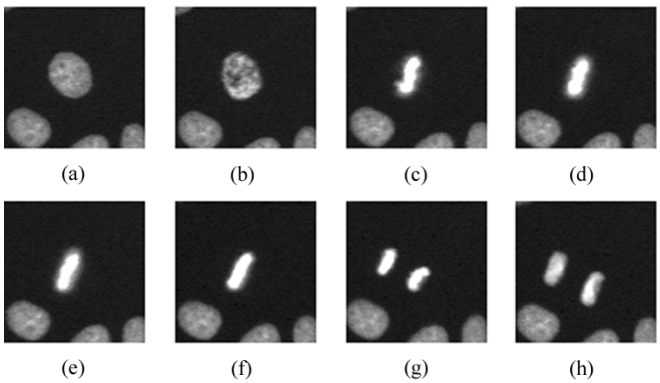

Fig. 1. Nucleus division and appearance changes during mitosis (a)-(h) consecutive image subframes taken from a time-lapse imaging sequence.

where C stands for compactness, P perimeter, and A area. C is 1 when the shape of a nucleus is circular. Compactness increases as the cell nucleus becomes less circular. If a round nucleus is divided into several fragments, the compactness of each fragment will be larger than the compactness of the entire nucleus. Metaphase and anaphase nuclei may have elongated shapes, which will result in a large value of compactness. If metaphase and anaphase nuclei are over-segmented, tiny fragments will be generated.

We developed a merging technique based on these observations. The procedure itself can be described as follows. Let N be the total number of segmented objects found by the watershed algorithm. Let T_{size} be the minimum size of a nucleus in the image. All touching objects are evaluated using a checking procedure. Two objects are considered touching if they belong to the same object before the application of the watershed algorithm. The merging process first finds the smallest touching object within each iteration and then uses the following checking process to update the segmentation till no more touching objects can be merged. The checking process is implemented as follows.

- If the size of an object is less than T_{size}, it is merged with its smallest touching neighbor. Isolated objects with sizes less than T_{size} are discarded.
- If the size of an object is greater than $T_{\mathrm{size},}$ three compactness values are calculated; namely, for the object, the object's smallest touching neighbor, and the two objects as a whole. If the merged object's compactness is the smallest, these two objects are merged.

B. Cell Phase Identification

Cell cycle progress can be divided into different stages (phases), such as interphase, prophase, metaphase, anaphase, and telophase. Cells under certain conditions may stay in metaphase and do not divide further, this condition is called "arrested" in metaphase. In our previous study [3], we extract a set of twelve features, i.e., maximum, minimum, and average of gray level, stand deviation of gray level, major axis, minor axis, elongation (major axis/minor axis), area (size of cell), perimeter, compactness, convex perimeter (perimeter of the convex hull), and roughness (convex perimeter/perimeter) to describe the variation of shape and intensity between different cell phases. We then run different combinations of feature selection, feature reduction, and classification tests. Experimental results show that a KNN ($K = 6$) classifier with a set of seven features generate the best classification results. In this paper, we expand the initial feature set with additional 28 co-occurrence features [5] and use the same KNN classifier for cell phase identification. Z-score, calculated by dividing a feature's distance to its mean by its standard deviation, is first employed to normalize the different features and the principal component analysis (PCA) is then employed to reduce the feature set. Here, Z-score is defined as $z_{ij} = (x_{ij} - \bar{m}_j)/s_j$, where x_{ij} represents $j-$ th feature of the ith nuclear and $\bar{m}_j$ is the mean value of all n cells for feature j, and s_j is the standard deviation of the jth feature. The detailed discussion is given in Section III-B.

A: 1,1,1,1,1,1,1,1,1,**2**,1,1,**2**,1,1,1,1,1,1,**2**,1,1,1,1,1,1,**2**,1,1,1,1,1,1,1

B: 1,1,1,1,1,1,1,1,1,**4**,1,1,1,1,**4**,1,1,1,1,1,1,**4**,1,1,1,1,1,1,1,1,1,1,1,1

C: **2**,**2**,**2**,**2**,**2**,**2**,**2**,5,5,5,5,5,5,5,5,5,5,5,**2**,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5

D: 1,1,1,1,1,1,1,1,2,**5**,3,3,3,3,3,3,4,4,4,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1

Fig. 2. Cases A-D show portions of cell sequences. 1 stands for interphase, 2 prophase, 3 metaphase, 4 anaphase or telophase, and 5 arrested metaphase.

Although cell cycle progress can be affected by adding drug treatments, the change is confined by biological constraints. Heuristic rules based on established biological knowledge can be applied to compensate certain phase identification errors. They are as follows.

Rule 1: The go-forward rule: Cell cycle progress can only go forward in the biologically-known cell phase sequence.

Rule 2: The continuation rule: Cell cycle progress cannot skip a cell phase and enter the phase after.

Rule 3: The phase-timing rule: The time period that a cell stays in a phase cannot be changed dramatically.

Fig. 2 provides four examples to illustrate the application of the heuristic rules to correct cell phase identification errors by the classifier. Each example in Fig. 2 indicates a small segment of nucleus sequences. Bold font marks the places where the errors happen. In Case A, the cell in the interphase is misclassified as a prophase cell four times: these errors can be detected and corrected by applying the go-forward rule. In Case B, the interphase cell is misclassified as an anaphase cell three times: these errors can be detected and corrected by applying the continuation rule. Cells in certain periods of prophase look similar to cells arrested in metaphase. This would cause certain misclassification between prophase cells and arrested metaphase cells. Fortunately, a cell can only stay in prophase for about 30 min, whereas for arrested metaphase forever. In Case C, the metaphase cell is misclassified as a prophase cell in the first seven frames, and these errors can be detected and corrected by applying the phase-timing rule. The remaining one error can be detected and corrected by applying the continuation rule. In Case D, one prophase cell is misclassified as an arrested metaphase cell. This error can be corrected by applying Rule 3, i.e., the phase timing rule.

C. Nuclei Tracking

After nucleus segmentation, a matching process is used to set up the correspondence between nuclei at time t and nuclei at time $t+1$ by computing the distances between them. An association matrix is used to measure the distance, which is defined as follows:

$$D = \begin{cases} 1 - \dfrac{C_t \cap C_{t+1}}{C_t \cup C_{t+1}} & C_t \cap C_{t+1} \neq 0 \\ 1 + \dfrac{D(C_t, C_{t+1})}{M(C_t, C_{t+1})} & C_t \cap C_{t+1} = 0 \end{cases} \tag{2}$$

where C stands for nucleus, $D(C_t, C_{t+1})$ is the Euclidean distance from the center of gravity of C_t to the center of gravity of C_{t+1}, and $M(C_t, C_{t+1})$ is the maximum size of C_t and C_{t+1}. A match is found if the distance is below a certain threshold. A large threshold is used during the matching processing which allows it to find all matching candidates. During the matching process, the following four cases can occur.

1) Only one nucleus at time $t+1$ matches a nucleus at time t.
2) No nucleus at $t+1$ matches a nucleus at t.
3) A nucleus at t matches more than one nucleus at $t+1$.
4) More than one nucleus at t matches a nucleus at $t+1$.

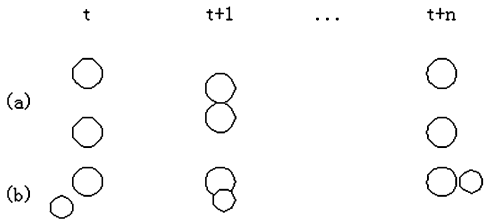

Fig. 3. Examples of ambiguous correspondence caused by under-segmentation: (a) Nuclei touching and (b) nuclei overlapping.

For Case 1, a successful match is found. Case 2 occurs when the nucleus moves out of the field of view. Only the border nuclei can move out of view. Thus, this situation can be detected by checking the nucleus position. Case 3 indicates a possible nucleus split. The fourth case shows possible under-segmentation. Case 3 and Case 4 cause ambiguous correspondences between the nuclei at t and $t+1$.

Case 3 and 4 can also be introduced by false matching, where no splitting or merging happens. We use the following procedure to remove false candidates founded by the matching processing. We first rank the candidates selected by the matching process according to the distances calculated: the smaller the distance the higher the rank. In a many-to-one case, we added the candidates' size together, highest rank first. We then compare the aggregated size of selected candidates with the size of the "one" nucleus. The true candidates are decided as these candidates whose aggregated size matched the 'one' nucleus most closely. In a one-to-many correspondence case, we use the same procedure to select true candidates.

We use the following strategy to resolve the issue of ambiguous correspondence. When nuclei cannot be separated from each other and ambiguous correspondence happens, information about these nuclei is recorded. Later, when they move away from each other, the tracking algorithm is able to solve the ambiguous correspondences by comparing their current status with previously stored information. In our method, nuclei size and their relative location, defined as a nucleus's location relative to another nucleus such as up, down, left, and right, are recorded. Fig. 3(a) illustrates the way that the tracking method solves ambiguous correspondence caused by nucleus touching. The position of top nucleus is always up to the position of the bottom nucleus. When the two touching nuclei move away from each other, their relative location does not change comparing with the time they move together. Thus, the relative location information can be used to solve such an ambiguous correspondence. Fig. 3(b) illustrates the way that the tracking method solves ambiguous correspondence caused by overlapping of nuclei. When the small nucleus moves away from the underneath big nucleus, correct correspondence can be set up by associating a nucleus to the one of similar size.

One kind of split is caused by over-segmentation in which a single nucleus is divided to multiple pieces due to incorrect segmentation. This kind of ambiguity can be identified by comparing the change of nucleus size and the relative location of each nucleus at time $t+1$. If the nuclei at $t+1$ are separated by one pixel with the sum of their sizes is the same as the nucleus at time t, the fragments are merged and viewed as a single unit.

Nucleus division can be considered as a special case of nucleus splitting where one nucleus divides into two or more daughter cell nuclei. If all daughter cell nuclei only correspond with one nucleus at t, these daughter cell nuclei are confirmed as the daughters of that nucleus. If multiple nuclei divide simultaneously together, ambiguous correspondences can happen. Fig. 4 shows an example of two nuclei dividing. In this example, the matching processing finds that daughter nucleus 4 corresponds with both nucleus 1 and 2, daughter cell nucleus 5 also

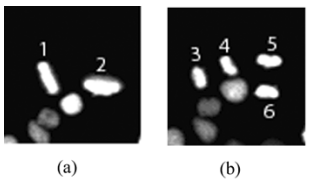

Fig. 4. An example of ambiguous correspondence caused by nucleus division: (a) Nuclei before division and (b) nuclei after division.

corresponds with both nucleus 1 and 2. To solve this ambiguous correspondence the center of gravity for every two daughter cell nuclei are calculated. For nucleus 4, we first calculate its center of gravity with nucleus 3, 5, and 6 separately. Then we calculated the distance from these three centers of gravity to each center of gravity for nucleus 1 and nucleus 2. Finally, we found the centers of gravity of nucleus 4 and 3 to the center of gravity of nucleus 1 is the smallest of them all, and thus nucleus 4 is determined to be the daughter cell nucleus of nucleus 1. Similarly, nucleus 5 is determined to the daughter cell nucleus of nucleus 2.

III. Evaluation Results and Discussion

Four nucleus sequences were used to test the efficiency of the proposed method. These sequences are generated by the Laboratory of Prof. R. King, Department of Cell Biology, Harvard Medical School. Each sequence consists of ninety-six frames over a period of 24 h. The sequences were recorded at a spatial resolution of 672×512, and a temporal resolution of one image per 15 min with a time-lapse fluorescence microscopy. Using a large time interval, we will not be able to capture certain prophase nuclei, as they only last for 15 min. On the other hand, exposure to high frequency of image sampling would damage and even kill the cells under observation. We use HeLa cells, one of the most widely used cell lines in biological research, in our experiments. Two types of sequences were used to denote drug treated and untreated cells. The drug used in our experiments is Taxol, an anti-mitotic drug. For the two drug-treated movies, the first one is scanned right after adding the drug while the second one is captured three days after adding the drug. Most of the cells in the second drug-treated movies were arrested in metaphase. The number of target nuclei in each sequence ranged from 78 to 204. After 24 h the number of nuclei could grow to more than 400 for untreated sequences. We developed a window-based C/C++ application program to implement the segmentation, classification and tracking algorithms presented in the paper. For an image with approximately 300 nuclei, the average computation time was 1.4 s on a Pentium IV 2.4-GHz computer. We discuss the evaluation results of nuclei segmentation, cell phase identification, and nuclei tracking in the following three subsections. All the results were compared with manual analysis results.

A. Segmentation

To test the segmentation algorithms, four images were chosen from each cell sequence. This generates a test set containing 16 images and 3 071 nuclei. The T_{size} used in the merging process were determined as 250 pixels, which is a low size threshold of nuclei and is decided experimentally. This value should be changed when the spatial resolution of the image changes. Fig. 5 shows several segmentation examples of these techniques. It can be seen clearly that the size and shape based merging technique can merge a lot of over-segmented nuclei that cannot be merged by the connectivity based techniques. It can be seen from Table I that 97.8% of the nuclei have been correctly segmented by watershed segmentation with shape and size based fragments merge.

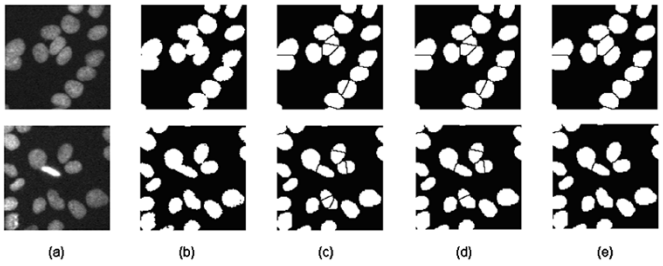

Fig. 5. Watershed segmentation results comparison. (a) Portion of original gray level images. (b) Binary images after ISODATA Thresholding. (c) Watershed segmentation results. (d) Watershed segmentation results with connectivity-based merging. (e) Watershed segmentation results with shape and size based merging.

TABLE I
NUCLEI SEGMENTATION RESULTS COMPARISON

	Nuclei Number	*Correct Separation*	*Over-Seg*	*Under-Seg*
Watershed	3,071	2,880 (93.8%)	165 (5.4%)	26 (0.8%)
Connectivity based merging		2,920 (95.1%)	105 (3.4%)	46 (1.5%)
Shape based merging		3,002 (97.8%)	29 (0.9%)	40 (1.3%)

To further evaluate the proposed scheme, we compare it with the segmentation results of the connectivity based merging [11], which merges nuclei fragments by checking the boundaries they shared with their neighbors. Table I shows the comparison results. The watershed algorithm caused 165 nuclei out of 3 071 nuclei to be over-segmented. The connectivity based technique can only merge 36.4% of them, while our new method merged 82.4%. The connectivity based merging technique failed because it was unable to deal with fragments whose size was larger than the preset value. In such cases, the fragments are considered as individual nuclei. The shape and size based technique causes 14 of the 2 880 separated nuclei to be merged, whereas the connectivity based technique causes 20.

In this study we use relative simple size and shape information to merge nuclei fragments which fail to merge certain nuclei. In our future study we will extract more information and explore statistical model methods to merge nuclei fragments.

B. Cell Phase Identification

A total of 200 nuclei were selected from the four sequences. Each nucleus was tracked for 12.5 h. Thus 50 images were taken for each nucleus. During this time, these 200 cells either divided or arrested in metaphase. This process generated a test set with 10 000 nuclei. The cell phase identification experiments were performed on this test set. The training set for the classifier consisted of the 400 nuclei used for feature reduction. Z-score is first employed to normalize the features and the same parameters are applied to test datasets. In PCA for feature reduction, we select the biggest eigenvalues that the sum of these eigenvalues is 90% of the sum of total eigenvalues. Eight features are obtained after using the PCA. We then test the 10 000 nuclei using KNN classifier with $K = 6$. Table II shows the results from automated analysis, which are compared with manual analysis results.

The classifier correctly identified nearly all (99%) interphase cells. For cells in metaphase and ana/telophase, the accuracy of the classifier algorithm was about 88% for both of them. However, only 68% of cells in prophase were correctly identified. The classifier made a number of mistakes on separating metaphase cells from prophase cells. 28% of prophase cells are assigned as metaphase cells, and 9.5% of metaphase cells are assigned as prophase cells. Table III lists the phase identification results achieved by the application of the knowledge-based heuristic rules to the classifier outputs. Note that most of the phase identification errors between prophase and metaphase have been corrected. Only 8% of prophase cells are assigned as metaphase cells, and 3.2% of metaphase cells are assigned as metaphase cells. As shown in Table III, the accuracy of phase identification increases by 16% for prophase cells, 7.8% for metaphase cells, 8.5% for ana/telophase cells, and 0.8% for interphase cells.

TABLE II
CELL PHASE IDENTIFICATION RESULTS (ASG: ASSIGNED, INTER: INTER PHASE, PRO: PROPHASE, META: METAPHASE, ANA/TELO: ANAPHASE/TELOPHASE, ACCU: ACCURACY, THE SAME AS TABLE III)

True \ Asg	*Inter*	*Pro*	*Meta*	*Ana/Telo*	*Accu (%)*
Inter (6832)	6764	3	58	5	99
Pro (125)	4	85	35	1	68
Meta (2523)	43	240	2225	14	88.2
Ana/Telo (520)	8	0	56	456	87.7

TABLE III
CELL PHASE IDENTIFICATION RESULTS BY APPLYING HEURISTIC RULES

True \ Asg	*Inter*	*Pro*	*Meta*	*Ana/Telo*	*Accu (%)*
Inter (6832)	6818	0	10	4	99.8
Pro (125)	8	105	10	2	84
Meta (2523)	14	80	2422	7	96
Ana/Telo (520)	4	0	16	500	96.2

In order to test how the number of training samples influence the recognition accuracy, we keet the same test data set with 10 000 nuclei from the 200 cell sequences, but reduced the training data to 100 nuclei. We randomly picked up the training samples from other 20 cell sequences which are not included in the 200 sequences. We found that interphase has 94.1%, prophase 73.4%, metaphase 90.7%, and anaphase 89.2% correct recognition accuracy. It is seen that the recognition accuracy of this case is lower than the case with training samples being 400 nuclei. That is reasonable as the training samples are dramatically reduced.

C. Tracking

To establish a metric for the performance of the tracking algorithm, three types of errors are considered below.

- *Percentage of nuclei tracked:* Number of nuclei tracked without termination through the entire sequence divided by the total number of nuclei in the beginning.
- *Percentage of divisions detected:* After nucleus division, daughter cell nuclei are detected and associated with their parents. This percentage is computed as the ratio of the number of divisions detected to the total number of divisions.
- *False division number:* This is the number of false divisions detected where two or more nuclei are associated with one nucleus in a previous frame which did not divide.

We test our tracking method on all four nuclei sequences. The images were taken over a few days. Nucleus which moves out of field of view during automated imaging is not counted in the final tracking statistic.

TABLE IV
NUCLEI TRACKING PERFORMANCE COMPARISON BETWEEN DIFFERENT TECHNIQUES (D: DETECTED, M: MISSED. THE SAME AS TABLE V)

Seq.	*Num*	*Loc.& Size Based Tracker*		*Centroid Tracker*	
		D	M	D	M
1	204	188	16	184	20
2	90	82	8	80	10
3	133	130	3	118	15
4	78	76	2	69	9
Total	505	476 94.3%	29 5.7%	451 89.3%	54 10.7%

TABLE V
NUCLEI DIVISION DETECTION RESULTS

Seq.	*Num*	*D*	*M*	*False Division*
1	62	57	5	6
2	57	51	6	5
3	80	79	1	0
4	0	0	0	3
Total	199	187 94%	12 6%	14

Table IV lists the tracking results. For each sequence, the tracking algorithm achieved more than 90% tracking accuracy.

We also compare our location-and-size based tracker with a centroid tracker [4]. The centroid tracker computes the "center of mass" within a subimage to locate the center of target. Here, "mass" is the sum of intensities in the subimage called the track gate. In this paper, we use the nuclei boundary founded by the Otsu's method as the boundary of track gate. The proposed tracking method achieved an average 94.3% tracking accuracy while the centroid tracker only achieves 89.3% tracking accuracy. The centroid tracker fails when nuclei touch or overlap each other and cannot be separated. By using the information about size and location of nuclei, our tracking method can successfully resolve the ambiguous correspondences caused by nucleus touching and overlapping and thus can used to increase the accuracy of cell tracking.

Table V provides the division detection performance of the proposed method. The tracking module correctly associates 94% daughter cell nuclei with their parents. Errors happen when daughter cell nuclei overlap with nearby nuclei right after division. In this case, the segmentation module cannot separate these daughters from the nuclei under them. This is the situation where most of the errors happen. False division is mostly caused by over-segmentation. Both situations can be handled by improving the efficiency of the nucleus segmentation module.

Yang *et al.* [13] also did similar work on dynamic cell analysis. They applied relatively simple segmentation and distance-based tracking methods to studying cell motility. However, the difficult problems in our time-lapse data sets are how to separate touching cells and how to solve the ambiguous correspondence. Our contribution here is that we systematically developed a series of advanced algorithms to address those two unresolved problems.

IV. CONCLUSION

Time-lapse fluorescence microscopy is an important tool in drug development and biological studies. Existing computational methods, however, are rather limited in dealing with large volumes of dynamic, image datasets generated from time-lapse fluorescence microscopy. In this paper, we have presented a new method to fill this gap. We developed a strategy of image analysis to resolve problems of touching cells and ambiguous correspondence and integrated them into a computational bioimaging system for automated segmentation, tracking and classification of cancer cell nuclei in time-lapse microscopy images. The cellular image analysis method can be used to segment, classify, and track individual cells in a dynamic cell population accurately and reliably. The experimental results reported in Section III for cancer cell cycle studies showed that the proposed method is efficient and effective in cell tracking and phase identification. The availability of such an automated analytic solution removes a major obstacle in the bioassay development using time-lapse florescence microscopy.

ACKNOWLEDGMENT

The authors would like to acknowledge the source of data and the collaboration with their biological science colleagues, and, in particular, Dr. R. King and Dr. S. Lyman, of the Department of Cell Biology, Harvard Medical School. They would also like to thank other members of the Life Science Imaging Group of the Center for Bioinformatics, Harvard Center for Neurodegeneration and Repair (HCNR) and Brigham and Women's Hospital, HMS for their technical comments. Finally, they would like to thank the reviewers for their constructive critique and suggestions in improving the quality of the presentation.

REFERENCES

[1] A. Bleau and J. L. Leon, "Watershed-based segmentation and region merging," *Comput. Vis. Image Understanding*, vol. 77, no. 3, pp. 317–370, 2000.
[2] X. Chen and S. T. C. Wong, "Automated dynamic cellular analysis in high throughput drug screens," presented at the IEEE Int. Symp. Circuits and System, Kobe, Japan, May 2005.
[3] X. Chen, X. Zhou, and S. T. C. Wong, "Knowledge-driven cell phase identification in time-lapse microscopy," presented at the IEEE Life Science Data Mining Workshop, Brighton, U.K., Nov. 2004.
[4] A. P. Goobic, M. E. Welser, S. T. Acton, and K. Ley, "Biomedical application of target tracking in clutter," in *Conf. Rec. Asilomar Conf. Signals, Systems and Computers*, vol. 1, 2001, pp. 88–92.
[5] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC–3, pp. 610–621, 1973.
[6] A. K. Jain, R. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, Jan. 2000.
[7] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contours models," *Int. J. Comput. Vis.*, vol. 1, pp. 321–331, 1988.
[8] G. Lin, U. Adiga, K. Olson, J. F. Guzowski, C. A. Barnes, and B. Roysam, "A hybrid 3D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks," *Cytometry A.*, vol. 56, no. 1, pp. 23–36, 2003.
[9] R. Nilanjan, S. T. Acton, and L. Klaus, "Tracking leukocytes *in vivo* with shape and size constrained active contours," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, pp. 1222–1235, Oct. 2002.
[10] N. Otsu, "A threshold selection method from gray level histogram," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, pp. 62–66, 1978.
[11] P. S. Umesh Adiga and B. B. Chaudhuri, "An efficient method based on watershed and rule-based merging for segmentation of 3-D histopathological images," *Pattern Recognit.*, vol. 34, no. 7, pp. 1449–1458, 2001.
[12] L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms," *IEEE Trans. Image Process.*, vol. 2, no. 2, pp. 176–201, Apr. 1993.
[13] F. Yang, G. Gallardo, M. A. Mackey, F. Ianzini, and M. Sonka, "Segmentation and quantitative analysis of the living tumor cells using large scale digital cell analysis system," Proc. SPIE, 2004, to be published.