Nuclei Segmentation Using Marker-Controlled Watershed, Tracking Using Mean-Shift, and Kalman Filter in Time-Lapse Microscopy

Xiaodong Yang, Houqiang Li, and Xiaobo Zhou

***Abstract*—It is important to observe and study cancer cells' cycle progression in order to better understand drug effects on cancer cells. Time-lapse microscopy imaging serves as an important method to measure the cycle progression of individual cells in a large population. Since manual analysis is unreasonably time consuming for the large volumes of time-lapse image data, automated image analysis is proposed. Existing approaches dealing with time-lapse image data are rather limited and often give inaccurate analysis results, especially in segmenting and tracking individual cells in a cell population. In this paper, we present a new approach to segment and track cell nuclei in time-lapse fluorescence image sequence. First, we propose a novel marker-controlled watershed based on mathematical morphology, which can effectively segment clustered cells with less oversegmentation. To further segment undersegmented cells or to merge oversegmented cells, context information among neighboring frames is employed, which is proved to be an effective strategy. Then, we design a tracking method based on modified mean shift algorithm, in which several kernels with adaptive scale, shape, and direction are designed. Finally, we combine mean-shift and Kalman filter to achieve a more robust cell nuclei tracking method than existing ones. Experimental results show that our method can obtain 98.8% segmentation accuracy, 97.4% cell division tracking accuracy, and 97.6% cell tracking accuracy.**

***Index Terms*—Cell segmentation, cell tracking, Kalman filter, mean shift, watershed.**

I. Introduction

ANALYZING cancer cells' cycle progression is important in understanding drug effects on cancer cells. Recent development of time-lapse microscopy provides an important tool to observe and measure the cell-cycle progression of individual cells in a large population. However, large volumes of image data generated by fluorescence microscopy make manual analysis time consuming and challenging. Therefore, automatic techniques to analyze cell cycles progress are of considerable interest in realizing the full potential of time-lapse microscopy in biological research or drug discovery [25], [26].

Manuscript received October 4, 2005; revised August 7, 2006. This work was supported by the Center for Bioinformatics, Harvard Center for Neurodegeneration and Repair (HCNR), Harvard Medical School. The work of Dr. X. Zhou was supported by HCNR Center for Bioinformatics Research Program and by the National Institutes of Health under Grant NIH R01 LM008696. This paper was recommended by Guest Editor P.-C. Chung.

X. Yang and H. Li are with the Department of Electrical Engineering and Information Science, University of Science and Technology of China, Hefei 230027, China (e-mail: xdyang@mail.ustc.edu.cn; lihq@ustc.edu.cn).

X. Zhou is with Center for Bioinformatics, Harvard Center for Neurodegeneration and Repair (HCNR), Harvard Medical School, Boston, MA 02115 USA, and also with the Department of Radiology, Brigham and Women's Hospital, Boston, MA 02115 USA (e-mail: zhou@crystal.harvard.edu).

Digital Object Identifier 10.1109/TCSI.2006.884469

A. Previous Works

A number of automated cellular image analysis methods have been proposed [1]–[7], most of which focus on segmentation and tracking problems of migrating cells in video-microscopy. Those methods will be summarized in the following paragraphs.

There are two approaches for automatic cell motility analysis [3]. One way is segmentation-based tracking strategy, which consists of two steps: segmentation and tracking. The objective of segmentation is to find cell candidates in a given frame. After segmentation, an association between frames is exploited on the detected cells. Such an approach makes use of boundary information, which has a good tracking accuracy. However, it is only efficient when object borders are sharp. The typical application of this method is shown in [1]. The other way is based on model adjustment to deal with cell motility. This method only focuses on tracking problem considering the fact that objects are generally difficult to be extracted. Other than segmenting objects, this method parameterizes and optimizes the shape of the cells so as to fit the cells to the targeted objects. In this way, the object trajectory can be traced from the frames. Snake [4]–[6], level set [7], and mean shift [3] are categorized in such methods. They have a common characteristic that the model parameters obtained from the former frame are used to initiate the process of the current frame. If objects have sharp borders and are relatively easy to be segmented, the segmentation-then-tracking approach is the first choice. Otherwise, the model adjustment approach is more suitable. It should be noticed that the model-based approach can also be used for objects with sharp borders. But it will cause a serious time-consuming problem, which would be even worse in the case of many objects being tracked.

Segmentation plays an important role in automatic cell motility analysis. If objects have sharp borders, simple thresholding method can work well enough to separate them from the background. Many existing methods contain such thresholding functions [8]. For example, ISODATA algorithm has been successfully applied to cell image segmentation [1], [9], in which the given gray scale image can be effectively and efficiently segmented. However, the main problem of this method is that it cannot separate touching objects. Watershed [8]–[11] is a valuable tool to tackle such problem. Original watershed directly uses region minimums or ultimate eroded points (UEPs) as starting points [1], [8]. Although it can segment touching

objects, oversegmentation would be likely to happen simultaneously. There are two approaches that can solve this problem: fragment merging and marker-controlled watershed. Two typical methods for fragment merging are connectivity-based merging method [12] and shape and size-based merging method [1]. Connectivity-based merging method [12] can only merge small fragment and will eventually fail if the size of fragment is above a certain preset value. Shape and size-based merging method [1] overcomes such shortcoming, but there are still much oversegmentation needing post-processing in the subsequent tracking step. Marker-controlled watershed [8] replaces region minimums or UEPs with predefined markers, each of which represents an object. Although marker-based method can effectively solve oversegmentation problem, it only works on the premise that the extracted markers really represent the true objects. Thus, the difficulty in performing this method is marker extraction. The proposed method in [9] successfully extracts cell makers and has a good result for cell segmentation, but this method only works for objects with descending intensity and round shape.

Comparing with segmentation, tracking is a more challenging job in cell motility analysis. Many efforts have been made to solve the tracking problem [1]–[7], [13]–[15]. Active contour techniques [16], [17] have been often used in cell tracking [4]–[6]. As mentioned above, it is a model adjustment approach, and can give a secondary result of segmentation of the objects. In this method, the tracking result of the former frame is used to initiate the parameters in the current frame and 100% tracking accuracy based on this method has been reported [4]. However, it requires cells to be partially overlapping in adjacent frames. Level set [7] is able to tackle object topology changes, but it will merge two contacting contours into a single one and it also requires cells to be partially overlapping in different frames. Another common drawback of snake and level set method is that they are iterative methods and thus time consuming, especially when many cells have to be tracked simultaneously. Recently, mean shift [18], [19] has been successfully applied to cell tracking [3]. Debeir *et al.* introduced the use of combination of linked kernels to track cell cytoplasm. Experimental results showed that Debeir's method was a good approach to track migrating cells. However, their method did not consider cell boundary detection, so they need post-processing to obtain boundary information. Centroid method [10] is fast and easy to deal with tracking. Its shortcoming is that satisfactory tracking performances could only be achieved when the inter-frame movement is relatively small comparing to cell size; otherwise, centroid tracker will fail. Chen *et al.* [1] have proposed a location-and-size-based tracking method to solve this problem, but it cannot track cell division effectively. Kalman filter has been applied to biological object tracking [13], but no explicit experimental result has been reported in the paper. However, it is still a good approach for predicting cell movement trajectory and it is a good assistant for mean shift [20], so we adopt it as a predicting tool in this paper.

B. Our Method

As discussed above, existing methods have limitations in dealing with microscopy datasets: first, many of them need manual or semi-manual initialization, which is inefficient in processing huge volume of images; second, cell segmentation has not been solved perfectly, especially for clustered cells; third, if cell movement is large, tracking will become very difficult; finally, these existing methods are time consuming.

To overcome the shortcoming of existing methods, we propose a new method combining marker-controlled watershed, Kalman filter and modified mean shift for cell motility analysis. The proposed method consists of following steps. Firstly, marker-controlled watershed is used to segment all cells, including touching ones. We do not employ UEPs as cell seeds, which often causes oversegmentation. Instead, in order to extract cell markers, we propose a group of cell-like masks to perform erosion, which is proven to be very robust and effective. Additionally, in order to improve segmentation accuracy, we utilize context information of video sequence. By extracting markers from the former frame or the next frame, oversegmentation and undersegmentation could be reduced to a low extent. Experimental result shows that our method can obtain high segmentation accuracy. Then, we track the cells in the full video sequence. Our tracking algorithm, which is based on mean shift and Kalman filter, could be performed on the full video sequence frame by frame. Because mean shift has the property of one-by-one tracking, we decide to track from the last frame to the first frame in order to detect cell division. Kalman filter employs the position of cell in the current frame to predict its corresponding position in the next frame, which enables our algorithm more robust. In addition, to make better use of shape and intensity information of cell nuclei, we modify classical Gaussian kernel (GK)-based mean shift. Three new kernels with scale, shape, and direction adaptation are proposed instead of Gaussian one due to the fact that, cell nuclei have ellipse shape in different directions and have nearly uniform intensity within the cell region. With such improvements, our modified mean shift algorithm is able to shift the position predicted from the former frame using Kalman filter to the target position which has largest kernel density in the current frame. We would like to point out that our method is different from [3], where the authors use weighted kernels for cell cytoplasm tracking. In our work, we invent new anisotropy and directed kernels for cell nuclei tracking. And our method is based on segmentation-then-tracking scheme, which can offers useful boundary information for classification purpose [1].

The rest of this paper is organized as follows. Section II introduces the watershed and our new marker extraction method. Section III gives mean shift tracking method combined with Kalman filter. Section IV presents the experimental results. Section V concludes this work.

II. Segmentation Using Marker-Controlled Watershed

In time-lapse fluorescence microscopy images, nuclei are bright objects within a dark background, which means that they can be easily segmented from background by histogram thresholding. However, thresholding method is not able to segment touching nuclei. Watershed is an effective tool to segment such cases, but classical watershed often leads to oversegmentation. Although existing marker-controlled watershed can solve oversegmentation, it cannot extract markers efficiently. In order to

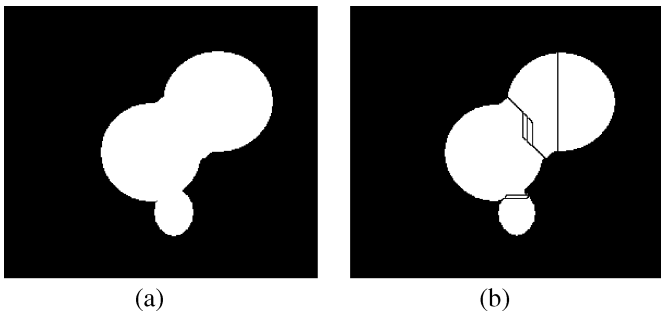


Fig. 1. Oversegmentation by classical watershed. (a) Binary image. (b) Segmented result by classical watershed.

segment clustered nuclei, we propose a new marker extraction method in this paper. In this section, we first briefly review the watershed algorithm and marker-controlled watershed. Then, the new marker extraction method is introduced and some experimental results are given. At last, we give the final segmentation results which have considered the information obtained from the former frame or the next frame.

A. Watershed and Marker-Controlled Watershed

The classical watershed [8] is based on flood simulation. Considering the input gray-scale image as a topographic surface, the segmentation problem is addressed to produce the watershed lines on this surface. Direct application of the watershed segmentation algorithm by flooding from the regional minimums or UEPs generally leads to oversegmentation due to noise and other local irregularities of the gradient. Too much oversegmentation will render the result of watershed algorithm useless, as shown in Fig. 1. There are two approaches capable of reducing this problem. One is to merge adjacent regions according to some criteria after the use of watershed. However, this approach is time consuming and merging criteria are very difficult to design. The other approach is marker-controlled watershed. A marker is a connected component of an image, which represents existence of an object. Marker-controlled watershed floods from the markers instead of flooding from the regional minimums or UEPs. If segmentation could be considered as a partition of an image, the binary marker image can be regarded as a first estimation of the partition. Then, the watershed grows the marker based on a flooding simulation process. Because one marker represents one object, it is important to extract the object marker correctly because improper operation on marker extraction will cause oversegmentation or undersegmentation. In the following section we propose a new marker extraction method to separate touching cells.

B. Our New Marker Extraction Method

In order to overcome the limitation of classical watershed, we present a new approach to deal with markers based on mathematical morphology. The proposed method is based on the concept of *condition erosion*: eroding only when the size of the object is larger than the predefined threshold. We have designed two series of masks for erosion operation according to the cell shape, which we refer to as *fine erosion structures* and *coarse erosion structures*, respectively. These masks can keep

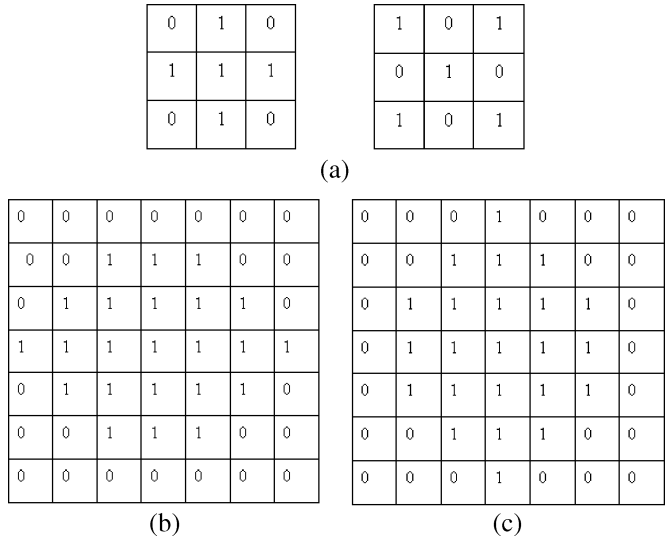


Fig. 2. Erosion structures. (a) Fine structures. (b) and (c) Two coarse structures.

the shape information in eroding operation, which is helpful for separating clustered cells. The *fine erosion structures* comprise two 3∗3 masks and *coarse erosion structures* have four types of 7∗7 masks, as shown in Fig. 2. These structures are designed according to the shape of the cell which is similar to an ellipse. The reason why we use two series of structures is that *coarse erosion structures* tend to remain the actual shape when reducing the size of clustered cells. However, they are likely to make the object disappear because of the dramatic reduction in size. On the other hand, *fine erosion structures* are less likely to make the object disappear, but they will lead to the loss of shape information, resulting in undersegmentation. So "first *coarse erosion structures*, and then *fine erosion structures*" is the proper procedure for cell nuclei segmentation.

To illustrate the shortcomings of coarse erosion structures and fine erosion structures, two examples are given in Fig. 3. Fig. 3 (a) and (b) shows two original binary images. Fig. 3 (c) gives the erosion result of Fig. 3 (a) by using coarse erosion structures. Only one object marker remains. Fig. 3 (d) shows erosion result of Fig. 3 (b) by using fine erosion structures. When only using fine erosion structures, shape information will be removed gradually, which results in incorrect extraction of object markers, as can be seen in Fig. 3 (d). Fig. 3 (e) and (f) shows the correct results by "first coarse, then fine" erosion strategy. Finally, Fig. 3 (g) and (h) gives watershed segmentation by our method. The fact that no oversegmentation exists proves the effectiveness of the proposed method.

The algorithm is described as follows.

1) Pre-processing: obtain the binary image, remove noise and predefine two thresholds T1 and T2 by experiment. T1 is the threshold of condition erosion with coarse erosion structures, and T2 is the threshold of condition erosion with fine erosion structures.
2) Erode iteratively on binary image with coarse erosion structures until the size of objects is smaller than T1.
3) Apply condition erosion with fine structures to the result obtained in step 2. The process is the same as step 2 except that the threshold is T2 instead of T1.

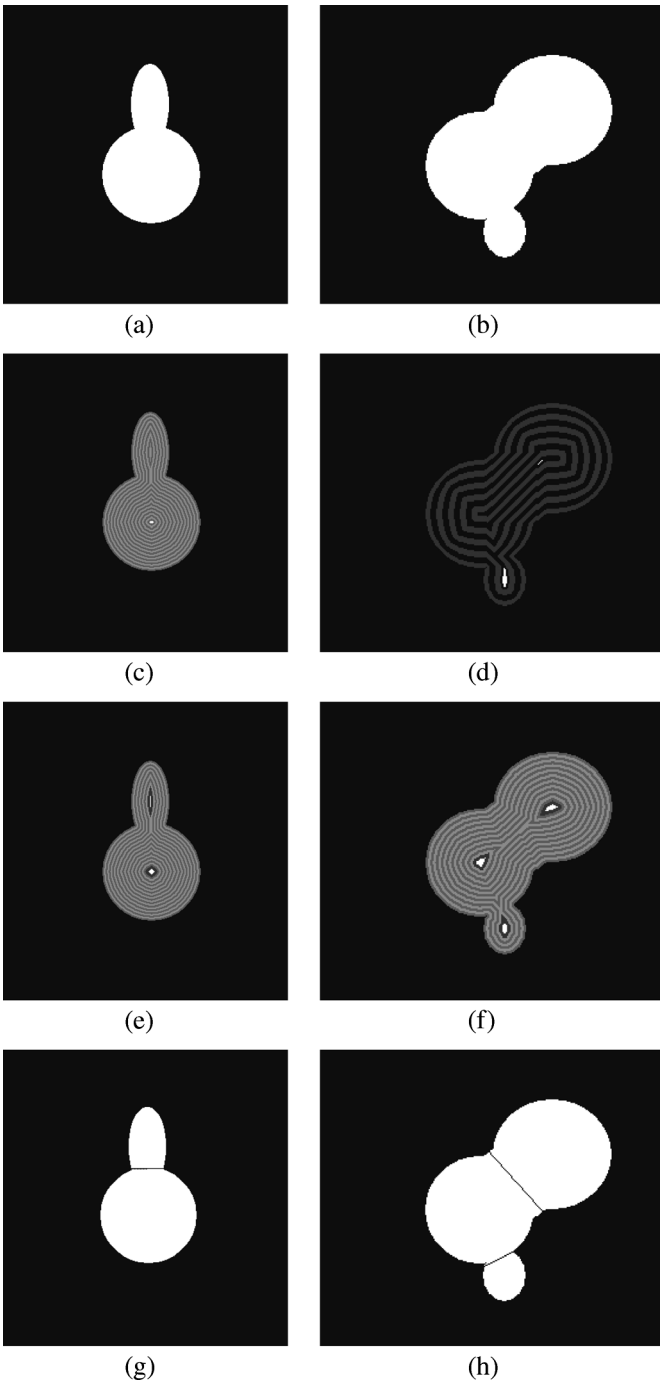


Fig. 3. Illustration of marker extraction. (a) and (b) are two original binary images. (c) Erosion result by coarse structures. (d) Erosion result by fine structures. (e) and (f) Markers extracted by first coarse then fine strategy. (g) and (h) Segmentation results by our method.

4) Extract the touching regions and their corresponding seed markers from the binary images obtained in step 1 and 3, and then calculate the euclidian distance map from the binary image made up of touching regions.
5) Apply marker-controlled watershed to euclidian distance map to segment clustered cells whose markers come from step 4.

In order to determine T1 and T2, we test a few cell nuclei images and compare the segmented results with those obtained by using classical watershed segmentation. It shows that T1 can be set as 200 and T2 can be set as 50. We then apply this method to other images. Fig. 4 shows some experimental results, where Fig. 4 (a) shows the original grayscale images. Fig. 4 (b) gives binary images. Fig. 4 (c) presents segmentation results by classical watershed, and Fig. 4 (d) shows segmentation results by our method. It shows that the proposed method seldom generates oversegmentation.

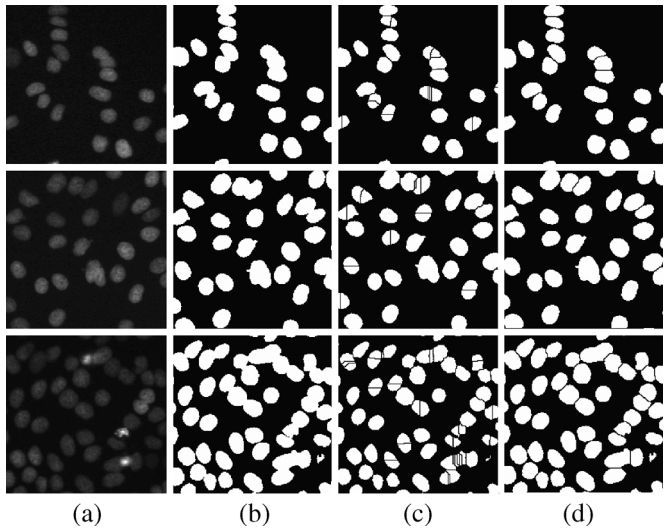


Fig. 4. Comparison of the proposed method with classical watershed. (a) Original gray scale image. (b) Binary images. (c) and (d) present segmentation results of two watershed methods.

Another advantage of our method is low time consumption. Unlike traditional watershed, our proposed method does not need to calculate all the Euclidian distances. The calculations of Euclidian distances only occur on the objects that have been considered as clustered regions; then, watershed segmentation on these regions is performed. Our method avoids the calculations of all the regions and saves large computation complexity.

C. Improve Over/Under Segmentation Using Context Information

Migrating of nucleus in time-lapse fluorescence microscopy is usually minute. Most cells do not move far away from their own positions in former frame. Because of migration, the degrees of cell overlapping may be different in consecutive frames. For example, cells in the current frame can be separated by marker-controlled watershed or they are already separated, but in next frame these touching cells cannot be separated by our watershed method because of high degree of overlapping. This inter-frame information can be utilized as an indication for segmentation of touching cells that can not be solved by the method proposed in Section II-B.

There are three cases we may meet, as shown in Fig. 5. First, cells are separated in the binary image of current frame but clustered in the next frame. Second, cells are separated after watershed in current frame and clustered in the next frame; the separation in current frame is caused by oversegmentation. Third, cells are separated by watershed in current frame and clustered in the next frame; separation in current frame is correct and clustering in the next frame is caused by undersegmentation. In the

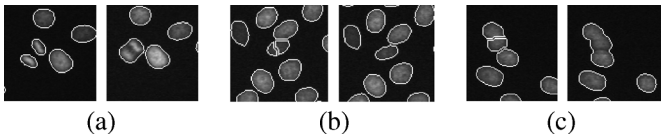


Fig. 5. Three cases needed to be corrected. (a) Cell in the right image should be separated with no doubt. (b) Oversegmentation cells in the left image should be merged. (c) Undersegmented cells in the right image should be separated with the help of shape and size-based merging technique [1].

first case, clustered cells in the next frame should be separated with no doubt. In the second case, oversegmented cells in current frame should be merged. In the third case, cells in the next frame should be separated. We need a criterion to judge whether to merge in current frame or to separate in the next frame. Fortunately, shape and size-based merging technique [1] provides a good approach to fulfill the judging task. In the first case and the third cases, clustered cells in the next frame are separated by marker-controlled watershed where the cell markers are extracted from current frame. It should be pointed out that the same strategy should be operated twice on the video sequence, forwardly and backwardly, respectively.

The modified segmentation algorithm is described as follows.

1) Segment all the images in the video sequence. Let t denote discrete time and set equal to 1.
2) Track from frame t to frame $t+1$. If two or more cells in frame t are tracked to one cell in frame $t+1$ and they are separated in the binary image of frame t, we can assert that the corresponding cell in frame $t+1$ consists of touching cells. The cell markers in frame t are adopted as markers in frame $t+1$ to segment the touching objects. If cells in frame t belong to the same cells in the binary image, use shape-and-size-based method to decide whether to merge in frame t or to separate in frame $t+1$.
3) Let $t = t+1$, then go to step 2 until t is equal to the number of the last frame.
4) Track from frame t to frame $t-1$. If two or more cells in frame t tracked to one cell in frame $t-1$ and they are separated in the binary image of frame t, we can deduce that the corresponding cell in frame $t-1$ consists of touching cells. The cell markers in frame t are adopted as markers in frame $t-1$ to segment the touching object. If cells in frame t belong to the same cell in the binary image, then use shape-and-size-based method to decide whether to merge in frame t or to separate in frame $t-1$.
5) Let $t = t-1$, then go to step 4 until t is equal to 1.

Fig. 6 provides some segmentation examples by using context information: Fig. 6 (a) and (b) shows segmented images in current frame and the next frame, respectively. Fig. 6 (c) gives the correct segmented results of the next images. From this figure, we can see that segmentation errors can be corrected by using context information.

III. Cell Tracking

Cell tracking is important in cell quantitative analysis; however, it is quite challenging to get satisfactory results when facing division cases. Luckily, mean shift [18], [19], which

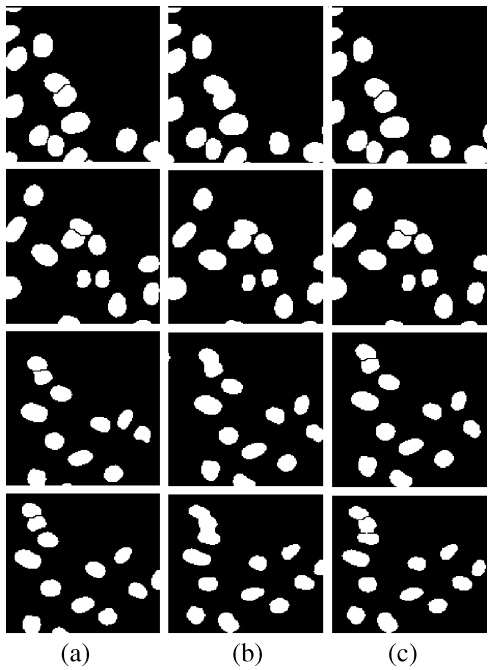


Fig. 6. Modified segmentation by adopting context information. (a) and (b) are segmented images in current frame and the next frame, respectively. (c) Correct segmented results of the next images.

bases on the notion of "kernel," is an efficient technique that can automatically sort out local model within a set of data. We propose to apply it to cell tracking in this study. In this section, the general mean shift theory will be first introduced; then, three novel adaptive kernels are developed to make the classical mean shift more robust for cell tracking; finally, the framework combining mean shift and Kalman filter is designed for cell tracking.

A. Mean Shift and Kernel Design

Let the kernel be G and $w(a)$ be image gray level in pixel a. The sample mean in x is

$$m(x) = \frac{\sum_a G(a-x)w(a)a}{\sum_a G(a-x)w(a)}. \quad (1)$$

The difference $m(x)-x$ is called the *mean shift*. The repeated movement of data points to the sample means is called the *mean shift algorithm*. The *mean shift algorithm* specifies how to combine the sample weights $w(a)$ in a local neighborhood with a set of kernel weights $G(a)$ to produce an offset that tracks the centroid of the objects in an image.

The most popular kernel in current papers [18], [19] is the GK

$$G(x,y,\sigma) = 1/(2\pi\sigma^2)\exp\left(-(x^2+y^2)/(2\sigma^2)\right). \quad (2)$$

Since this kernel has only one parameter, it becomes impractical when object shape is changing. We design another three kernels considering the cell shape and gray level information.

1) Oriented Ellipse Kernel: Due to the reason that the shape of an object may be changing as a function of time and be radially asymmetric sometimes, we have to design other kernels to make the mean shift algorithm more robust. A new kernel is proposed to replace the GK. We refer to this kernel as "oriented

ellipse kernel (EK)," because its shape is very similar to ellipse. The EK is

$$G(x,y,a,b,\theta) = \frac{1}{2\pi ab} \times \exp\left(-\left(\frac{(x\cos\theta + y\sin\theta)^2}{2a^2} + \frac{(x\sin\theta - y\cos\theta)^2}{2b^2}\right)\right). \quad (3)$$

The parameters a and b determine the compactness of ellipse. Parameter θ determines the direction of ellipse. We can adjust these parameters according to the shape of objects.

2) Non-Gaussian Kernel: All the kernels used in previous methods are based on Gaussian function, assuming that the gray-scale of the object follows Gaussian distribution; however, this is not always the truth. In fact, borders of objects may have great discontinuity against background and the grayscale inside the object is nearly uniform. These cases make GK impractical for object tracking. Therefore, we propose a non-Gaussian kernel (NGK) to solve this problem. The kernel here is very similar to low-pass filter, which consists of two following parts

$$G(x,y,t,\sigma) = \begin{cases} 1, & (x^2+y^2)^{1/2} < t \\ \exp\left(-(x^2+y^2-t)^2/(2\sigma^2)\right), & (x^2+y^2)^{1/2} \geq t \end{cases} \quad (4)$$

where t is related to the object scale and σ indicates the descending rate from 1 to 0. In this kernel, two parameters make the kernel more controllable for object tracking and we simplify t as σ.

3) Non-Gaussian Ellipse Kernel: The NGK can be combined with EK. We call this kernel as the "non-Gaussian ellipse kernel (NGEK)." It has properties of both EK and NGK.

B. Parameter Selection

After we propose several kernels with different parameters, how to set these parameters becomes a problem.

1) Scale Selection: To make the algorithm practical, the ellipse parameters a and b are represented by s *and* σ

$$\begin{cases} a = \sigma * s \\ b = \sigma / s \end{cases} \quad (5)$$

where s is called compactness and σ is named scale.

The four kernels mentioned above share the same parameter, namely kernel size. When the kernel size is too large, the tracking window will contain many background pixels as well as the foreground object pixels. If the size is too small, it will lead to poor object localization. In the worst case, tracking error will happen due to the small kernel size. Collins [21] used scale space theory to automatically select the kernel size. Scale space was first introduced by Witkin [22]; later, Lindeberg [23] developed the theory, making it an elegant theory to select the best scale for describing the features in an image.

Let the Gaussian filter be (2), and Laplacian of Gaussian (LOG) filter be

$$\mathrm{LOG}(x,\sigma) = \frac{2\sigma^2 - \|x\|^2}{2\pi\sigma^6} e^{-\frac{\|x\|^2}{2\sigma^2}}. \quad (6)$$

The convolution of image $f(x)$ with a LOG operator of scale σ is defined as

$$L(x,\sigma) = \sum_a \mathrm{LOG}(a-x,\sigma) f(a). \quad (7)$$

Lindeberg [23] showed that $L(x,\sigma) = s^2 L(sx, s\sigma)$, which means that the LOG operator is not invariant to scale. However, $\sigma^2\mathrm{LOG}(x,\sigma)$ is a normalized operator whose magnitude stays invariant during changes of image scale. Thus, in our application, we prefer to compute the convolution of image with $\sigma^2\mathrm{LOG}(x,\sigma)$ and select σ maximizing convolution results as the cell scale, which is used as the mean shift tracking scale.

The algorithm proposed in [21] is very time consuming, because the algorithm tries to make the location and scale converge at the same time. We find that the scale doesn't change dramatically between consecutive frames in our experiment. Therefore, the scale in the former frame can be considered as the approximate kernel size in the current frame. When the location in the current frame has been tracked, the scale is updated using scale space with the new detected location. This procedure is repeated until all frames have been tracked.

2) Direction and Compactness Selection: Direction is an important parameter for the anisotropy EKs. It has been observed that the direction of cell changes gradually. First we use the direction determined in the former frame in mean shift algorithm. After the location has been tracked and the scale has been updated, direction is updated according to updated location and scale. With the updated scale and location, kernel density is computed in different directions at this position. The direction, which maximizes the kernel density based on the updated location and scale, is selected as the kernel direction in the current frame. Similar to direction selection, the one that maximizes kernel density is selected as the updated compactness s.

C. Kernel Selection

Four kernels are introduced in Section III-B, including one popular GK and three newly proposed kernels. These four kernels are the GK, the EK, the NGK, and the NGEK. The first and the third ones are isotropy kernels, while the second and the fourth ones are anisotropy kernels. In order to select an ideal kernel for cell tracking, a quantitative criterion is required to evaluate the performance of each kernel. We use information of both gray scale image and binary image to design such a criterion. The mean shift result obtained in gray scale image is defined as estimated centroid, which belongs to a region in the corresponding binary image. We compute the centroid of this region, and defined this centroid as the actual centroid. Mean shift error is defined as absolute difference of these two centroids. In general conclusion, we average the errors from some test cells. The smaller the error, the more robust the kernel is.

In cell tracking, cell size is changing during time. Scale adaptation is required to make mean shift robust. Since scale is very important for tracking size-varying objects, all kernels are tested with scale adaptation. With several randomly selected cells, the errors of four kernels are computed, which are 1.9 for the GK, 1.5 are EK, 1.6 for NGK and 1.3 for NGEK. The NGK is selected due to its smallest mean shift error of all the four kernels.

D. Kalman Filter

One problem in the mean shift tracking is that sometimes the target cell is not within the target region centered at the last object position. This situation can be attributed to three reasons. The first one is that the object is moving too fast. The second

one is the low frame rate. More memory is needed to record huge volume images and more time is consumed to process the data if the frame rate is too high. Lastly, the search region is too small. In order to prevent this problem, the search region could be adjusted to be larger; another problem of including other objects in the search region would arise at the same time. Luckily, this problem could be resolved by Kalman filter method.

Kalman filter is an optimal recursive data processing algorithm [13], which combines all available measurement data, plus prior knowledge about the system and measuring devices, to produce an estimate of the desired variables in such a manner that the error is minimized statistically. The Kalman filter consists of two steps: prediction and correction. In the former step, the state is predicted with the dynamic model, while in the latter one, it is corrected with the observation model. Since the error covariance of the estimator is minimized, it can be regarded as an optimal estimator.

We denote the variable as: the discrete time:t, state vector: $X(t)$, observation vector: $Z(t)$, state transition matrix: A, observation matrix: C, state noise: $\nu(t)$, observation noise: $\mu(t)$, and covariance matrix: $P(t)$. The system is expressed as

$$\begin{cases} X(t+1) = AX(t) + \nu(t) \\ Z(t) = CX(t) + \mu(t). \end{cases} \tag{8}$$

Here we assume $\nu(t)$ and $\mu(t)$ are Gaussian random variable with zero mean, so their probability distributions are $p(\nu) \sim N(0, Q)$, $p(\mu) \sim N(0, R)$, where the covariance matrix Q and R are referred to as transition noise covariance matrix and observation noise covariance matrix. More information of Kalman filter can be referenced to [24].

We design a model to track migrating cells, the detail is as follows. The state vector is $X = [x, y, v_x, v_y, a_x, a_y]^T$, where (x, y), (v_x, v_y) and (a_x, a_y) represent position, velocity and acceleration, respectively. The observation vector is $Z = (x, y)$. The state transition matrix is

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0.5 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and the observation matrix is

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

E. Combined Mean Shift and Kalman Filter Tracking

As discussed above, we adopt the "NGEK" in mean shift tracking process. Because the mean shift technique only has the ability of one by one tracking, a feasible solution of tracking the cell division is to track from the last frame to the first frame.

The mean shift is computed using (1). Note that G represents the NGEK here, while sample weight $w(a)$ directly uses the input image. Tracking process of one cell is described as follows. First, the initial position of cell in the last frame is established and the initial width is given by the area of the cell. Then, following the backward direction, shift the position established in the former frame to a new position in the current frame with the width also established in the former frame. At last, update the position with mean shift tracking result. After the position has been established, direction, scale and compactness are then updated. This procedure is operated iteratively till the first frame. The combined mean-shift and Kalman filter tracking algorithm is summarized as follows.

First, select the cells to be tracked in the last frame. The selection procedure is processed automatically. Only the area size of cell within a range (50∼600 in our experiment) is selected as candidate cell. The reason is that large object may consist of clustered cells and small object is often caused by noise. Cells near the boundary are also discarded because these cells have high probability to move out of view, which will cause tracking errors.

Second, initialize parameters. Transition noise covariance matrix Q and observation noise covariance matrix R should be initialized before tracking. Several cells are selected for statistical purpose. Track these cells without using Kalman filter and make sure that the tracking results are correct. Then observation errors and prediction errors are calculated based on the tracking results without Kalman filter. Q and R's initialization is completed by computing covariance matrixes of observation errors and prediction errors. Other parameters needed to be initialized in Kalman filter are also estimated in this procedure. The location parameter is set to be the centroid of each candidate cell. Because we don't have any *a priori* information about the movement of objects in a sequence before tracking, we simply set the initial velocity and acceleration parameters to **zero**.

Third, predict state vector and state error covariance matrix.

Fourth, get the observation vector. The observation vectors are obtained by searching around the region centered at the predicted position. This process consists of two steps. If the predicted position is within one cell in the current frame, the centroid of this cell is considered as observation vector. Else, mean shift is used to search the cell centroid in the grayscale image. The searching result is served as the observation vector.

Fifth, correct the state vector and state error covariance.

Sixth, iterate the procedure from third to fifth until it reaches the first frame.

In the end, the flow chart of our tracking system is shown in Fig. 7.

IV. Experimental Results

To verify the efficiency of the proposed algorithm, we carried out a series of experiments on the microscopy video sequence, which consists of 240 frames over a period of 60h. The sequence was recorded at a spatial resolution of 672 × 512, and the temporal resolution of one image per 15 minutes with a time-lapse fluorescence. The cell type is HeLa cell, one of the most widely used cell lines in biological research. Our goal is to segment all the cells and track each of them from the last frame to the first frame correctly.

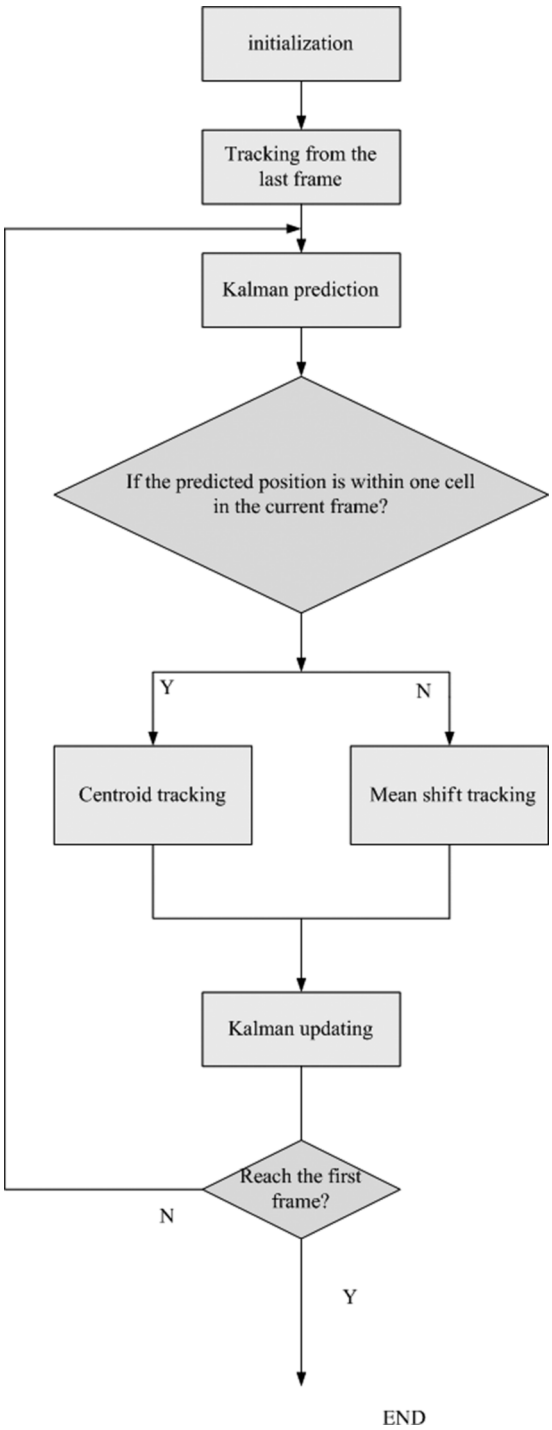


Fig. 7. Proposed tracking system.

A. Segmentation

We randomly selected five images from the sequence with 1178 cells to validate the proposed segmentation algorithm. Fig. 4 shows some segmentation examples of the proposed algorithm, from which it can be seen clearly that our method can segment most of touching cells with little oversegmentation and undersegmentation. To demonstrate the improvement of our method, we compared our method with simple watershed algorithm [8], and watershed with shape and size-based merging [1]. Table I shows the comparison of segmentation result with different techniques. The method proposed here correctly segments 98.8% of the cell nuclei. Classical watershed can correctly segment 93.8% of the nuclei. Shape and size-based merging technique can obtain 97.8% accuracy of nuclei segmentation. It is clear that the proposed method is superior to the other two methods.

Although the proposed marker-controlled watershed is superior to the conventional watershed, in the experiment we found that there were still some clustered cells that could not be correctly separated. In order to improve the segmentation performance, the proposed context-based method is applied. From Fig. 6, it can be seen clearly that this method can segment most touching cells, which cannot be segmented with only using the marker-controlled watershed. 0.5% undersegmentation and 0.2% oversegmentation have been reduced. The improvement on segmentation is very valuable because it has an important and direct impact on tracking.

B. Tracking

As discussed above, the NGEK is the best kernel for mean shift in our cell tracking application among four kernels. So we adopt it in mean shift and applied this mean shift to our proposed tracking framework combined with Kalman filter.

Two parameters are important to evaluate the efficiency of system performance in a cell tracking system. One is cell tracking accuracy and the other is division detection accuracy. Our method is compared with Centroid tracker [15] and location and size-based tracker [1]. Table II gives comparison results of tracking performance. We selected 82 cells to be tracked in the sequence. 2 cells were missed. The rest, 80 cells were successfully tracked. The tracking accuracy is 97.6%, while centroid tracker's is 89.3% and location and size-based tracker's is 94.3%. As for cell division, our method can track 97.4% cell division, while location and size-based Tracker can only track 94% cell division. In comparison, we conclude that our method can track cell migration and detect cell division more effectively.

V. Conclusion

This paper presents an automatic method to segment and track cells in time-lapse microscopy sequences, which enables analyzing cell cycle progress to be easier and more robust. Comparing with existing methods, our proposed method is fully automatic. Furthermore, it is able to segment and track migrating cells with higher performance, which offers a better understand of drug effects on cancer cells. The proposed method is based on watershed and mean shift. A new marker-controlled watershed is introduced to segment clustered cells efficiently. We propose different kernels with more freedom, such as scale, direction, and compactness, to adapt the cells' shapes, which makes the mean shift tracking process more robust. We also propose the combined Kalman filter and mean shift tracking. At last, we test our method on a microscopy video sequence. The experimental results show that the proposed method can obtain high segmentation and tracking accuracy.

TABLE I
SEGMENTATION COMPARISON OF DIFFERENT TECHNIQUES

	Cell number	Correctly segmented	Over-segmented	Under-segmented
Watershed	1178	93.8%	5.4%	0.8%
Shape and size based merging		97.8%	0.9%	1.3%
Our method		1164 (98.8%)	6 (0.5%)	8 (0.7%)

TABLE II
COMPARISON OF CELL TRACKING PERFORMANCES

	Centroid Tracker	Location and Size based Tracker	Kalman and Mean Shift
Tracked	89.3%	94.3%	97.6% (80/82)
Missed	10.7%	5.7%	2.4% (2/82)

ACKNOWLEDGMENT

The authors would like to acknowledge the excellent collaboration with their biology collaborators in this research effort, and, in particular, Dr. R. King of the Department of Cell Biology and ICCB (Institute of Chemistry and Cell Biology), Harvard Medical School. They would like to thank Dr. S. T. C. Wong, the Director of the Center for Bioinformatics, Harvard Center for Neurodegeneration and Repair (HCNR) and Brigham and Women's Hospital, HMS for his technical comments.

REFERENCES

[1] X. Chen, X. Zhou, and S. T. C. Wong, "Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 4, pp. 762–766, Apr. 2006.

[2] X. Yang, H. Li, X. Zhou, and S. T. C. Wong, "Automated segmentation and tracking of cells in time-lapse microscopy using watershed and mean shift," in *Proc. 2005 Int. Symp. Intell. Signal Process. Commun. Syst. (ISPACS'05)*, Dec. 2005, pp. 533–536.

[3] O. Debeir, P. V. Ham, R. Kiss, and C. Decaestecker, "Tracking of migrating cells under phase-contrast video microscopy with combined mean-shift processes," *IEEE Trans. Med. Imag.*, vol. 24, no. 6, pp. 697–711, Jun. 2005.

[4] Z. Christophe, L. Elisabeth, M. Y. Vannary, G. Nancy, C. Jean, and M. Olivo, "Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: A tool for cell-based drug testing," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, pp. 1212–1221, Oct. 2002.

[5] N. Ray, S. T. Acton, and K. Ley, "Tracking leukocytes *in vivo* with shape and size constrained active contours," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, pp. 1222–1235, Oct. 2002.

[6] Meas-Yedid and J. C. Olivo_Marin, "Active contours for the movement and motility analysis of biological objects," in *Proc. IEEE Int. Conf. Image Process.*, 2000, vol. 1, pp. 196–199.

[7] D. P. Mukherjee, N. Ray, and S. T. Acton, "Level set analysis for leukocyte detection and tracking," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 562–572, Apr. 2004.

[8] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Upper Saddle River, NJ: Prentice-Hall, 2002.

[9] N. Malpica, C. Ortiz de Solórzano, 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*, no. 28, pp. 289–297, 1997.

[10] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 6, pp. 583–598, Jun. 1991.

[11] A. Bleau and L. J. Leon, "Watershed-based segmentation and region merging," *Comput. Vision Image Understanding*, vol. 77, no. 3, pp. 317–370, 2000.

[12] P. S. U. Adiga and B. B. Chaudhuri, "An efficient method based on watershed and rule-based merging for segmentation of 3-D histo-pathological images," *Pattern Recogn.*, vol. 34, pp. 1449–1458, Jul. 2001.

[13] A. Genovesio and J.-C. Olivo-Marin, "Tracking fluorescent spots in biological video microscopy," in *Proc. SPIE Three-Dimensional Multidimensional Microscopy: Image Acquisition and Processing X*, 2003, vol. 4964, pp. 98–105.

[14] Z. Wang, B. Hu, L. Liang, and Q. Ji, "Cell detection and tracking for micromanipulation vision system of cell-operation robot," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, 2000, vol. 3, pp. 1592–1597.

[15] 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, Syst. Comput.*, 2001, vol. 1, pp. 88–92.

[16] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contours models," *Int. J. Comput. Vision*, no. 1, pp. 321–331, 1988.

[17] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 359–369, Mar. 1998.

[18] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, no. 8, pp. 790–799, Aug. 1995.

[19] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.

[20] D. Comaniciu and V. Ramesh, "Mean shift and optimal prediction for efficient object tracking," in *Proc. Int. Conf. Image Process.*, 2000, pp. 70–73.

[21] R. T. Collins, "Mean-shift blob tracking through scale space," in *Proc. CVPR*, 2003, pp. 234–241.

[22] A. P. Witkin, "Scale space filtering," in *Proc. Int. Joint Conf. Artificial Intell.*, 1983, pp. 1019–1022.

[23] T. Lindeberg, "Feature detection with automatic scale selection," *Int. J. comput. Vision*, vol. 30, no. 2, pp. 79–116, Nov. 1998.

[24] G. Welch and G. Bishop, "An introduction to the Kalman filter," in *Proc. ACM SIGGRAPH Conf.*, Los Angeles, CA, Aug. 2001.

[25] X. Zhou and S. T. C. Wong, "High content cellular imaging for drug development," *IEEE Signal Process. Mag.*, vol. 23, no. 2, pp. 63–72, Feb. 2006.

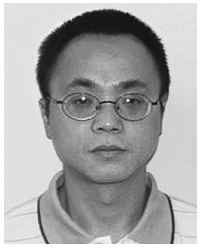
[26] ——, "Informatics challenges of high throughput microscopy," *IEEE Signal Process. Mag.*, vol. 23, no. 3, pp. 170–174, Mar. 2006.

Xiaodong Yang received the B.S. and M.S. degrees in electrical engineering from University of Science and Technology of China (USTC), Hefei, China, in 2003 and 2006, respectively.

His research interests include image processing, computer vision, pattern recognition, and wavelet analysis.

Houqiang Li received the B.S., M.S., and Ph.D. degrees in electrical engineering from University of Science and Technology of China (USTC), Hefei, China, in 1992, 1997, and 2000, respectively.

He is currently an Associate Professor in the Department of Electrical Engineering and Information Science, USTC. His research interests include image and video processing, video coding, computer vision, pattern recognition and bioinformatics.



Xiaobo Zhou received the B.S. degree in mathematics from Lanzhou University, Lanzhou, China, in 1988, and the M.S. and Ph.D. degrees in mathematics from Peking University, Beijing, China, in 1995 and 1998, respectively.

From 1988 to 1992, he was a Lecturer at the Training Center, 18th Building Company, Chongqing, China. From 1992 to 1998, he was a Research Assistant and Teaching Assistant in the Department of Mathematics, Peking University, Beijing, China. From 1998 to 1999, he was a Postdoctoral Research Fellow in the Department of Automation, Tsinghua University, Beijing, China. From 1999 to 2000, he was a Senior Technical Manager with the 3G Wireless Communication Department, Huawei Technologies Co., Ltd., Beijing. From February 2000 to December 2000, he was a Postdoctoral Research Fellow in the Department of Computer Science, University of Missouri-Columbia, Columbia, MO. From 2001 to 2003, he was a Postdoctoral Research Fellow in the Department of Electrical Engineering, Texas A&M University, College Station. From 2003 to 2005, he was a Research Scientist with the Harvard Center for Neurodegeneration and Repair, Harvard Medical School, Boston, MA, and also with the Radiology Department, Brigham and Women's Hospital, Boston, MA. Since 2005, he has been an Instructor in the Radiology Department, Brigham and Women's Hospital and Harvard Medical School. His current research interests include image and signal processing for high-content molecular and cellular imaging analysis, informatics for integrated multiscale and multimodality Bio/Medical imaging analysis, molecular imaging informatics, neuroinformatics, bioinformatics for genomics and proteomics. He has published more than 100 papers and holds five patents. He has authored one book titled *Computational System Bioinformatics*, which is to be published.

Dr. Zhou is a Member of SPIE, and Sigma Xi and he was listed in *Who's Who in Science and Engineering* in 2004.