15th European Signal Processing Conference (EUSIPCO 2007), Poznan, Poland, September 3-7, 2007, copyright by EURASIP

AUTOMATIC APPROACH FOR ANALYZING THE ENDOSOME MOVEMENT IN FLUORESCENCE VIDEO MICROSCOPY

Carol Rus, Mikko Vänninen, and Ulla Ruotsalainen

Institute of Signal Processing, Tampere University of Technology, Finland
email: {carol.rus, ulla.ruotsalainen}@tut.fi, mikko.vanninen@gmail.com
web: http://sp.cs.tut.fi/

ABSTRACT

Endosomes are structures located in a biological cell and their movement can be interesting for studying the effect of pharmaceutical substances on cell processes. The movement of endosomes can be analyzed by using fluorescence microscopy and by producing videos of few seconds length. Movement analysis done by a human specialist is usually unappealing because it is both time consuming and difficult, also the results depend always on the person doing the analysis. By automating the movement analysis we can ease and speed up the process, and the results would also be easily replicable. The paper presents a new tracking approach for endosome movement based on the double exponential smoothing. When compared to Kalman filter, a commonly used movement predictor, the proposed method proved to be simpler to implement, with lower running time and with good results. Also, the paper shows ways to express the movement of the tracked endosomes.

1. INTRODUCTION

An endosome is a structure (also called organelle) located in a biological cell that is involved in the transport of proteins from the outside to the inside of the cell. When looking to a cell's endosomes, they move in some direction with a particular speed and they can change their movement characteristics, when some substances are administered to the body that cell belongs to. As an application example, in this way, knowledge can be gained about the effects that pharmaceutical substances have on cell processes. The movement of endosomes can be studied using fluorescent microscopy and producing videos of few seconds length. The videos are produced in a digital format, i.e. AVI. The movement in these videos can be analyzed manually by a specialized person. This is a difficult and time consuming task, since the endosomes seem to move completely randomly at a first glance. An automatic approach for this problem can give the results faster and in a more objective way.

Tracking endosome movement deals with ambiguous situations, like partial or complete overlapping and focus issues. There have been proposed some methods to solve this problem; these are usually either deterministic or stochastic. Sage et al. [3] uses dynamic programming to track a single particle, even if it dims over time or is absent completely. This approach assumes that the movement is only a few pixels from frame to frame. This does not seem to be suited very well for our problem that concerns with multiple objects that are fast moving. Han et al. [4] adopts neural networks for tracking multiple objects. This approach is said to be able to reliably handle irregular motion, occlusions and changing appearances. It does this by postponing the decision until it has enough information for it, by keeping multiple hypotheses of trajectories in a graph structure. This algorithm promises to give a solution for the problem we want to solve, but the complexity is high. Also, it was developed for security cameras. Main problem arises from the fact that people look unique, endosomes do not, so that tracking may lead to difficulties to make a distinction between two possibilities. The *Tensor Voting* method was proposed by Kornprobst et al. [5] for noisy sequences. It uses a perceptual grouping, coding the estimated velocity, number of neighbors involved and their coherency. This method is computational intensive, and it may present problems when trying to track similar looking objects, especially when objects disappear and then later reappear.

Genovesio et al. [6] developed a solution based on using a Kalman prediction method together with an association technique for tracking fluorescent spots. This method is able to deal with ambiguity and varying situations, like objects size changes. *Double exponential smoothing* approach was proposed by LaViola [7] for object tracking in virtual reality. The method proved to give at least as good results as Kalman filter. It is simpler to implement and the running time is lower. This paper tries to solve the tracking problem for endosome movement analysis in fluorescence microscopy by using a method based on *double exponential smoothing*, and it is organized as follows. The next section gives details about the approach proposed. Section 3 gives experimental results and discussion, and the last section draws some conclusions and tells about future work.

2. METHOD

In fluorescent microscopy, the endosomes are seen as green spots that have no clear border. They can change their direction and speed at any given time. The fluorescence creates uneven labeling for the objects. We can only see one focal plane in the image sequence, and as the microscope has a limited focus depth some endosomes may disappear and then reappear in the focus. Also, since the endosomes are moving in a 3D space, they can obscure one another, either partially or completely at some instance in time.

The proposed method for tracking the endosomes comprises of four stages: image enhancement, segmentation, estimation and association. Figure 1 summarizes the key elements and the processing flow of the approach discussed by the paper.

For image enhancement step, we use first a low-pass filtering to attenuate the noise, then we enhance the endosomes edges with a high-boost filtering [1] and the last phase is the contrast-limited histogram equalization [2] in order to enhance the contrast within each frame of the AVI file.

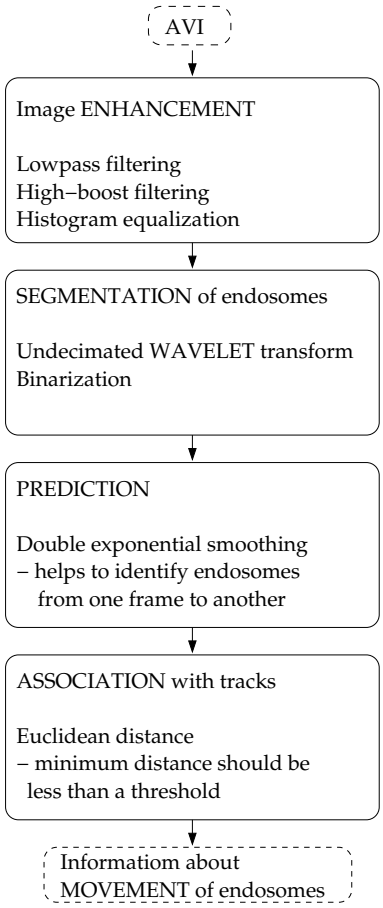


Figure 1: Diagram of the approach proposed for the automatic movement analysis of endosomes.

Segmentation is performed using an *undecimated wavelet transform* method [8]. The wavelet transform is computed row by row, followed by column by column, convolution with the b-spline kernel $[1/16;1/4;3/8;1/4;1/16]$. The discontinuity problems at the image borders are solved by symmetric mirroring extension. The process can be repeated recursively and it can be described by the next equations:

$$W_1 = A_0 - A_1 \tag{1a}$$
$$W_i = A_{i-1} - A_i \tag{1b}$$

where W_i are images containing the wavelet coefficients, and A_i are recursively obtained approximations of the original image A_0 by convolution with the above mentioned kernel, $i = 1,\ldots,J$. The number of iterations (also, called *scales*) is indicated by J. The term *undecimated* comes because this wavelet transform does not subsample by 2 the result after each iteration, as a normal wavelet transform does. The next step in segmentation is to threshold the wavelet coefficients. The threshold value is scale dependent and it is computed based on median absolute deviation. Equation 2 shows how the threshold value was computed in our implementation.

$$thr_i = 5 \cdot median|W_i - median(W_i)|, \tag{2}$$

where *median* indicates the midpoint in a series of numbers, and W_i represents a series of all wavelet coefficients at scale i. Wavelet coefficients that have a value greater than thr_i keep their values and the rest of them are set to zero.

The last step of segmentation is to obtain a binary image that indicates the endosomes. First, a correlation (across scales) matrix is computed by simply doing the element-wise product (also, called Hadamard product) of W_i matrices. Second, the computed correlation matrix is thresholded. In our implementation we set to 1 (i.e. endosome) only values very close (i.e. 0.5% deviation) to the maximum value in the correlation matrix. An example result can be seen in Figure 2.

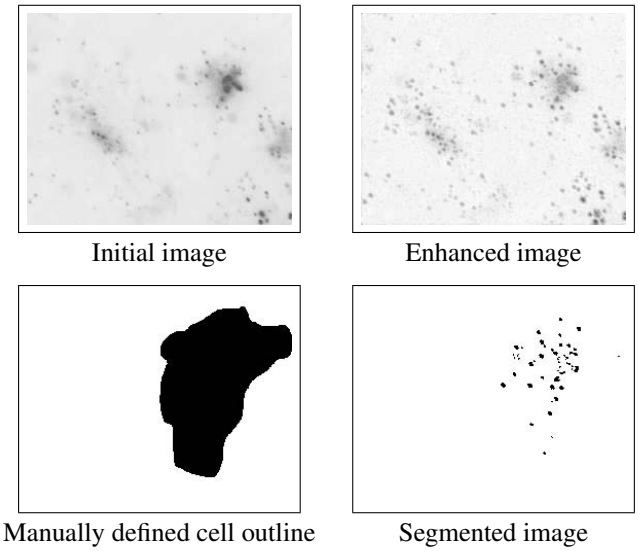


Figure 2: A segmented frame: initial image, enhanced image, the mask that shows the cell of interest within the frame and the final result of endosomes segmentation.

After we have found the objects in the image we try to associate them to a previous track or in case we cannot, we create a new track. This association is done by simply calculating the Euclidean distance between the estimates (predictions) given based on the previous frames and the centroids of the spots segmented in the current frame. For an estimate, if there is no distance less than a threshold value between centroids in the current frame and the estimate then the track is considered terminated.

For all the tracks, new and continuing, the prediction of the next positions is done by using a double exponential smoothing approach.

The tracks are stored in structure of size MxN, where N is the total number of tracks detected and M is the total number of frames in the video file. Each element in the structure is a position vector that keeps an x and y value.

A double exponential smoothing model gives to the past observations exponentially smaller weights as the observations get older. In other words, recent observations are given relatively more weight in prediction than the older observations.

There are two equations associated with *double exponen-*

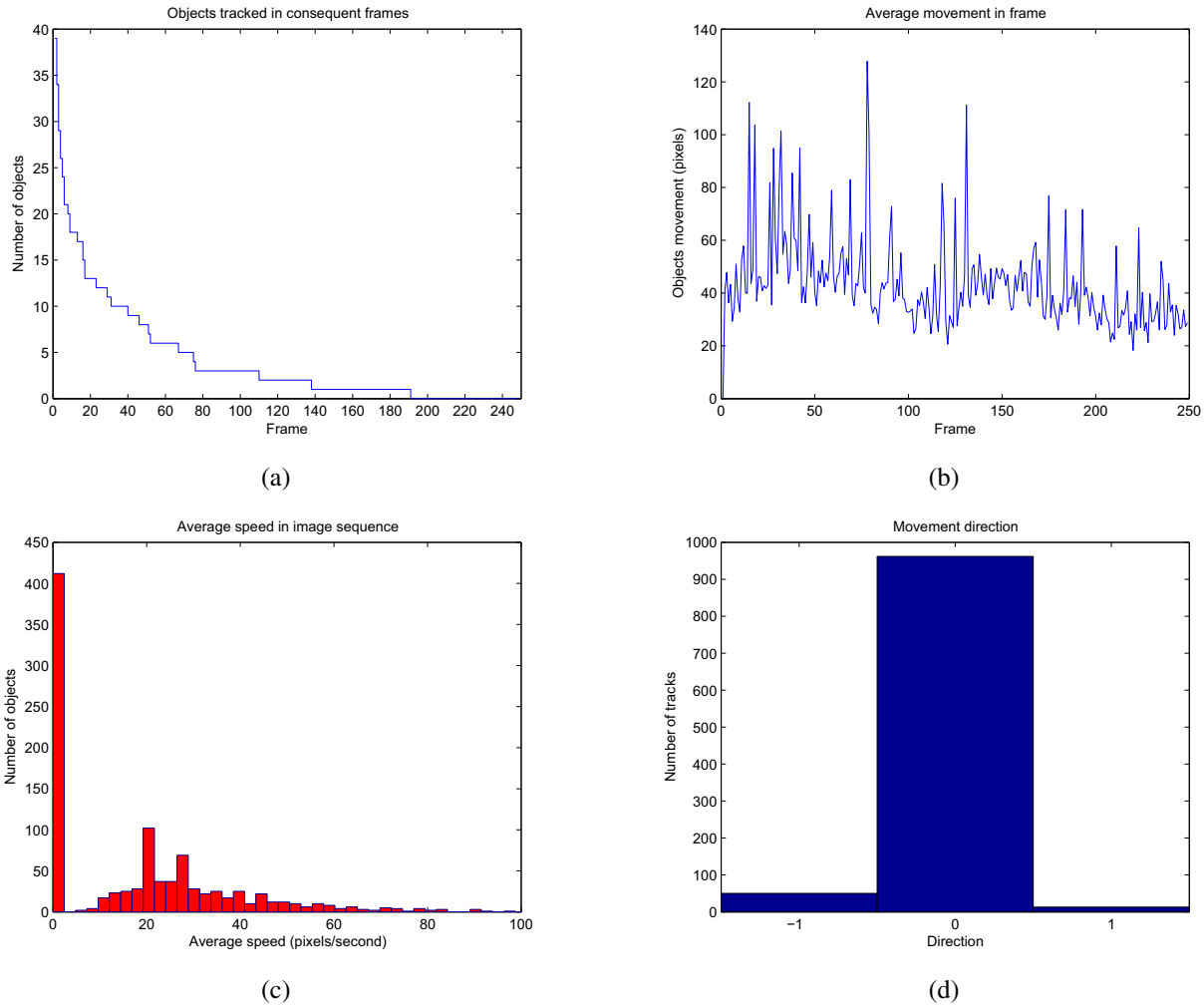


Figure 3: Movement quantization for cell MLN_02 (see Figure 4). (a) Number of object tracked in each frame; (b) Average movement (expressed in pixels) in each frame; (c) Histogram of the average speed for the tracks detected; (d) Histogram of the direction quantization for the tracks detected: -1 denotes a movement towards the cell nucleus centroid, 0 shows a movement around the nucleus centroid or no significant movement, and $+1$ indicates a movement towards cell membrane.

tial smoothing (shortly, DESP), as follows:

$$f_t = a \cdot Y_t + (1-a) \cdot (f_{t-1} + b_{t-1}) \quad (3a)$$
$$b_t = g \cdot (f_t - f_{t-1}) + (1-g) \cdot b_{t-1} \quad (3b)$$

where Y_t is the observed value at time t, f_t is the prediction at time t, b_t is the estimated slope at time t, a is the first smoothing constant, used to smooth the observations, and g is the second smoothing constant, used to smooth the tendency of the data.

The DESP estimation is given by equation:

$$f_{t+1} = f_t + b_t \quad (4)$$

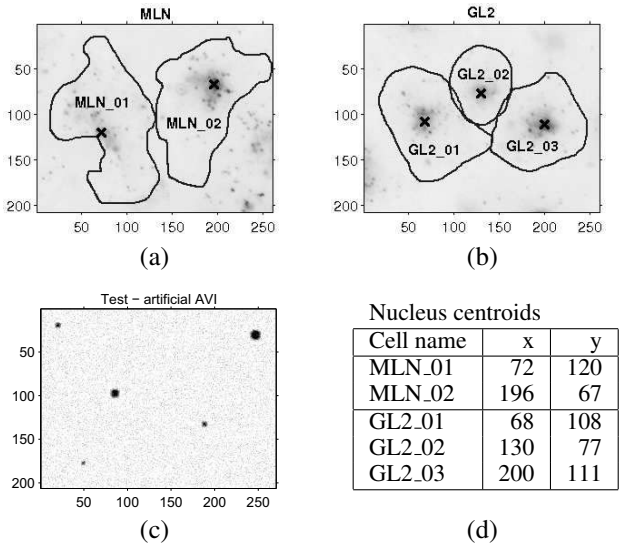
To initialize the double exponential smoothing model, f_1 is set to Y_1, and the initial slope b_1 is set to the difference between the first two observations (i.e. $Y_2 - Y_1$). The smoothing constants have values in the range $0.0 - -1.0$. When a smoothing constant is close to 1.0, more weight is given to recent observations, and when it is close to 0.0, less weight is given to recent observations.

When applied for movement prediction in a video file, the first frame that can be estimated by a double exponential smoothing predictor is frame number 3, considering that the counting starts at 1. For initialization purposes, the first frame of the video is doubled. This was considered since in a real microscopy avi the number of segmented endosomes might differ between frame 1 and frame 2, and those are important for starting the double exponential smoothing algorithm.

Based on the tracks detected, a quantization of the endosome movement can be given. Two useful movement indicators are the speed and, respectively, the direction. The *speed* is quantized by computing the average speed for each track and then by displaying the average speed distribution as a histogram. For *direction* an auxiliary parameter is needed: *the centroid of the cell nucleus*. The direction is quantized using three values, as follows: -1 denotes a movement towards the cell nucleus centroid, 0 shows a movement around the nu-

cleus centroid or no significant movement, and $+1$ indicates a movement towards cell membrane. Let us consider d_{intial} the distance between the initial point of the track and the nucleus centroid. And d_{final} the distance between the final point of the track and the nucleus centroid. Then, if $d_{final} - d_{intial}$ is positive the direction is coded as $+1$, if $d_{final} - d_{intial}$ is negative the direction is quantized as -1, and else, if the difference is zero the movement is coded as 0. Figure 3 (c) and (d) shows a movement quantization of speed and, respectively, direction, for one of the tested cells.

3. RESULTS AND DISCUSSION



Nucleus centroids

Cell name	x	y
MLN_01	72	120
MLN_02	196	67
GL2_01	68	108
GL2_02	130	77
GL2_03	200	111

(d)

Figure 4: Cells used for testing the approach. (a) MLN video file containing two cells; (b) GL2 AVI file containing three cells; (c) AVI file created artificially for testing the tracking algorithms; (d) A table that shows the centroid position for the real cells. Also, the centroids are marked by an "x" on the images.

The proposed approach was tested using real video AVI files produced using fluorescent confocal microscopy. Each frame within the AVI file was an image of 270x206 pixels and the frame rate was 20 frames per second. Images were a little noisy and blurry. The tested AVI files had on average 250 frames. Each AVI file contained in focus more than one cell. The cell were separated by using a mask outlined by a human specialist, which, also, pointed out the coordinates of the cell nucleus centroid. Beside the real video files, an artificial video file was created to test the tracking performance for which the centroid was not considered. Figure 4 shows all the cells that were tested, and their centroid.

Figure 5 presents a tracking result using double exponential smoothing prediction by showing different time instances. It reveals some of the difficulties of the problem: endosome overlapping since the movement is in 3D, and hard conditions for segmentation, since very crowded (also, some of them overlapped) spots may appear from time to time. Also, the figure reveals that the number of endosomes tracked decreases when the number of frame increases; situation showed by the Figure 3 (a), too. The point (b) of the same figure is usefull to see movement peaks and valleys across the frames.

In the implementation, we used $a = 0.2$ and $b = 0.3$ for the constants within the double exponential smoothing equations, and we considered a threshold value equals with 40 for the association step.

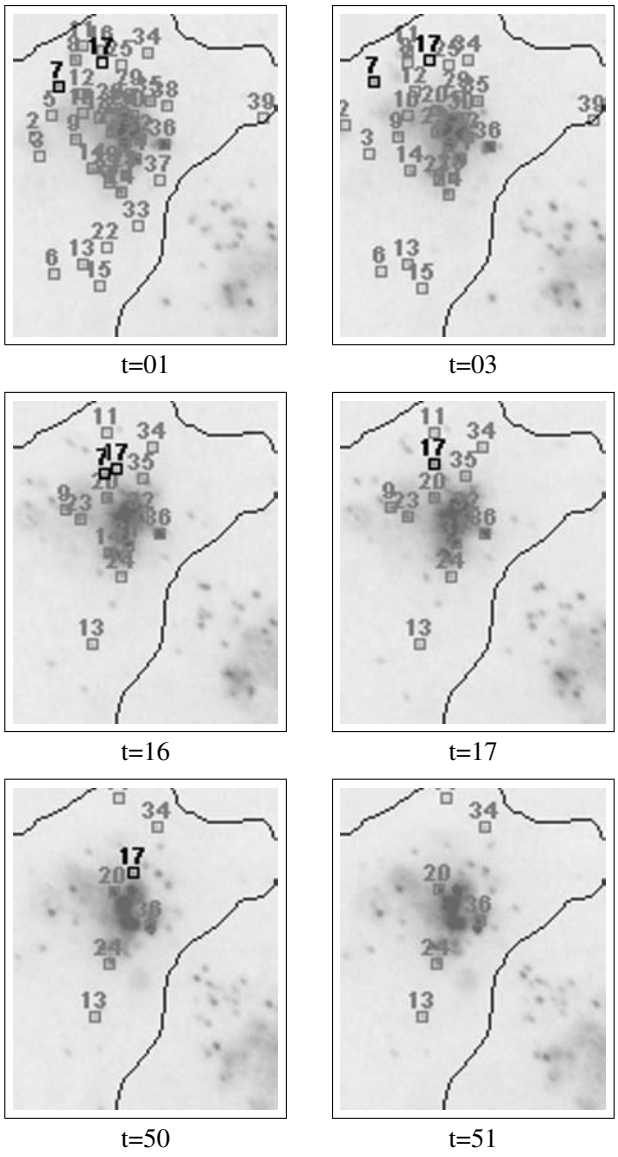


Figure 5: Tracking result for cell MLN_02. The images have been cropped to make the presentation clearer. The black line shows the cell border. Objects(endosomes) labeled 7 and 17 are outlined. At frame $t = 17$ they become overlapped, since the movement is in a 3D space, and track 7 is terminated. At frame $t = 51$ due to segmentation limitation track 17 ends.

When compared to Kalman filter (see Table 1), the approach was faster with up to 10 seconds on the tested data, and the quality of the results was good. It can be seen also that the number of tracks given by DESP tends to be a bit larger than the number of tracks given by Kalman filter. Applied on the test artificial video data, both algorithms lost a track when two objects came too close to each other.

The proposed method has lower complexity than a

Cell	Measurements	DESP	Kalman
MLN_02	Processing time (sec)	105	115
	Total number of objects detected	8236	8236
	Max number of objects in one frame	46	46
	Min number of objects in one frame	24	24
	Average number of objects in one frame	33.07	33.07
	Total number of TRACKS detected	1025	1012
	Average track length (given in frames)	8.03	8.13
GL2_03	Processing time (sec)	102	109
	Total number of objects detected	8551	8551
	Max number of objects in one frame	45	45
	Min number of objects in one frame	29	29
	Average number of objects in one frame	36.08	36.08
	Total number of TRACKS detected	1038	1026
	Average track length (given in frames)	8.23	8.33
Test data - artificially created	Processing time (sec)	8.4	8.5
	Total number of objects detected	170	170
	Max number of objects in one frame	6	6
	Min number of objects in one frame	4	4
	Average number of objects in one frame	5	5
	Total number of TRACKS detected	8	7
	Average track length (given in frames)	21.25	24.28
	Real number of objects	5	5

Table 1: Double exponential smoothing vs. Kalman filter. Some of the measurements presented had the same value for both cases, because they depended on the segmentation algorithm which was the same for both cases, and they were presented to reveal the performance of the chosen segmentation approach.

Kalman filter, while giving the same good results.

4. CONCLUSION

The paper presented a new tracking approach for endosome movement based on the double exponential smoothing and some methods to quantize the endosome movement. The method is able to follow all objects present in a frame, but due to losing objects in the segmentation stage either to not being able to detect them at all or to objects obscuring one another, the average track length is about 8 frames. Improving on the image enhancement or/and estimation can give better tracking result.

The wavelet based segmentation works quite well. As a future work, we will try to find better solutions for the image enhancement step, and to optimize the a and b constants within double exponential smoothing equations using an algorithm as Levenberg-Marquard.

REFERENCES

[1] R.C. Gonzalez, and R.E. Woods, "Digital Image Processing," Addison-Wesley publishing Co., 2nd edition, 2002.

[2] Karel Zuiderveld, "Contrast Limited Adaptive Histogram Equalization," Graphics Gems IV, p. 474-485, code: p. 479-484, 1994.

[3] D. Sage, F. Neumann, F. Hediger, S. Gasser, and M. Unser, "Automatic tracking of individual fluorescence particles: Application to the study of chromosome dynamics," *IEEE Trans. Image Processing*, vol. 14, no. 9, pp. 1372–1383, Sept. 2005.

[4] M. Han, A. Sethi, W. Hua, and Y. Gong, "A detection-based multiple object tracking method," in *Proc. IEEE Int. Conf. on Image Processing, ICIP'04*, vol. 5, Oct. 2004, pp. 3065–3068.

[5] P. Kornprobst and G. Medioni, "Tracking segmented objects using tensor voting," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*, vol. 2, June 2000, pp. 118–125.

[6] A. Genovesio and J.-C. Olivo-Martin, "Tracking fluorescent spots in biological video microscopy," in *Proc. SPIE Int. Conf. on Three-Dimensional and Multidimensional Microscopy*, 2003.

[7] Joseph J. LaViola Jr., "Double exponential smoothing: An alternative to kalman filter-based predictive tracking," in *Proc. of Eurographics Workshop on Virtual Environments*, 2003.

[8] J.-C. Olivo-Martin, "Extraction of spots in biological images using multiscale products," *Pattern Recognition*, vol. 35, pp. 1989–1996, 2002.