Automatic Cell Tracking of Microscopic Imaging of Living Cells by a novel Similarity-based Tracking Algorithm, and the Description of the Kinetics of the Cells for Image-Mining

Petra Perner[1]

[1] Institute of Computer Vision and applied Computer Sciences (IBaI), Arno-Nitzsche-Str. 45, 04251 Leipzig, Germany

Live-cell assays are used to study the dynamic functional cellular processes in High-Content Screening (HCA) of drug discovery processes or in computational biology experiments. The large amount of image data created during the screening requires automatic image-analysis procedures that can describe these dynamic processes. One class of tasks in this application is the tracking of cells. We describe in this paper a fast and robust cell tracking algorithm applied to High-Content Screening in drug discovery or computational biology experiments. We developed a similarity-based tracking algorithm that can track the cells without an initialization phase of the parameters of the tracker. The similarity-based detection algorithm is robust enough to find similar cells although small changes in the cell morphology have been occurred. The cell tracking algorithm can track normal cells as well as mitotic cells by classifying the cells based on our previously developed texture classifier. Results for the cell path are given on a test series from a real drug discovery process. We present the path of the cell and the low-level features that describe the path of the cell. This information can be used for further image mining of high-level descriptions of the kinetics of the cells.

Keywords: Cell Tracking; Similarity-Based Cell Tracking; Microscopic Image Analysis; Description of the Kinetic of the Cells; High-Content Analysis; Computational Cell Biology; Image Mining

1. Introduction

The utilization of dynamic High-Content Analysis approaches during preclinical drug research will permit to gain a more specified and detailed insight into complex sub-cellular processes by the use of living cell culture systems. This will effectively support future drug discoveries leading to a great outcome of highly specific and most effective drugs that come along with a well improved compliance. Live-cell assays are therefore used to study the dynamic functional cellular processes in High-Content Screening of drug-discovery processes or in computational biology experiments. The large amount of image data created during the screening requires automatic image analysis procedures that can automatically describe these dynamic processes. One class of tasks in this application is the tracking of the cells, the description of the events and the changes in the cell characteristics, so that the desired information can be extracted from it based on data-mining and knowledge-discovery methods.

There are several tracking approaches known that track a cell by detection in each single frame and associate the detected cells in each frame by optimizing a certain probabilistic function [1]. Other approaches track cells by model evolution [2]. This approach seems to be able to handle touching and overlapping cells well but is computational expensive [3].

We propose in this paper a similarity-based approach for motion detection of cells that is robust and fast enough to be used in real world applications. The algorithm is based on a specific similarity measure that can detect cells although small changes in morphology appeared. The algorithm can track normal cells as well as mitotic cells. In Section 2 we review related work. The material is given in Section 3 and the basic constraints of the algorithm resulting from microscopic life-cell images are explained. The algorithm is described in Section 4. Results are given on a test series from a real drug-discovery process in Section 5. The features that describe the extracted path are explained in Section 6. Finally, we give conclusions in Section 7.

2. Related Work

Li et al. [4] proposes a system that combines bottom-up and top-down image analysis by integrating multiple collaborative modules, which exploit a fast geometric active contour tracker in conjunction with adaptive interacting multiple models motion filtering and spatiotemporal trajectory optimization. The system needs ten frames to initialize its parameters. If the number of frames in a video is high this is acceptable otherwise it is not. Cells that have not been recognized in the initial frame are not considered anymore. The system cannot track new appearing cells and it has a high computation time. The evaluation of the categorical analysis shows that it has a high rate of false detection of mitosis and new cells. Swapping of cell identities occurred mostly in densely populated regions, where the boundaries between cells are highly blurred. This poses the question why to use snakes for tracking the cells.

In Sacan et al. [5], a software package is described for cell tracking and motility analysis based on active contours. Since snakes relies on the assumption that the movement and deformation of the tracked object is small between

consecutive frames they propose like an ensemble of different tracking methods in a coarse to refined fashion in order to make their method more robust. Their combined method performs the following steps: the overall displacement and rotation of the object is first determined using a template matching method. The resulting contour is used as the initial state to pyramidal Lucas–Kanade optical flow-based deformation; statistical outlier detection and local interpolation is applied to achieve resistance against errors in the optical flow evaluation. They report that the ensemble method achieves accurate tracking even for large displacements or deformations of the objects between frames but a deep evaluation of the method is not given in the paper. The speed, area, deformation, trajectory and detailed tracking of the cells are computed and displayed for analysis. Besides automated cell detection and tracking capability, the interface also allows manual editing to initialize or modify the tracking data. The software can be used to work with movie or image files of a variety of file formats. The tracking results can be exported either as raw text data for further numerical analysis, or as movie or image files for visualization, sharing and publishing. Cell lineage is not reported in the paper.

In Wang et al. [6] is reported a system that can segment, track, and classify cells into cell types such as interphase, prophase, metaphase and anaphase. They use the distance transform to binarize the image and seeded watershed transformation to detect the cell area. The tracking path is determined by using local tree matching method. Graph matching approaches require large computation time. For each cell are calculated a large set of features such as texture, shape, gray level. The relevant feature set is determined by using a feature selection procedure. On-line support-vector machine is used to classify the nuclei. This approach is able to adapt the classifier to changing appearances of the objects.

Cohen et al. [7] propose an algorithmic information theoretic method for object-level summarization of meaningful changes in image sequences. Object extraction and tracking data are represented as an attributed tracking graph (ATG), whose connected subgraphs are compared using an adaptive information distance measure, aided by a closed-form multi-dimensional quantization. The summary is the clustering result and feature subset that maximize the gap statistic.

In Maška et. al [10] the coherence-enhancing diffusion filtering is applied on each frame to reduce the amount of noise and enhance flow-like structures. Then, the cell boundaries are detected by minimizing the Chan–Vese model in the fast level set-like and graph cut frameworks. To allow simultaneous tracking of multiple cells over time, both frameworks have been integrated with a topological prior exploiting the object indication function.

Magnusson et. al [11] propose a global track linking algorithm, which links cell outlines generated by a segmentation algorithm into tracks. The algorithm adds tracks to the image sequence one at a time, in a way which uses information from the complete image sequence in every linking decision. This is achieved by finding the tracks which give the largest possible increases to a probabilistically motivated scoring function, using the Viterbi algorithm.

Padfield et. al [12] present a general, consistent, and extensible tracking approach that explicitly models cell behaviors in a graph-theoretic framework. They introduce a way of extending the standard minimum-cost flow algorithm to account for mitosis and merging events through a coupling operation on particular edges. They show how the resulting graph can be efficiently solved using algorithms such as linear programming to choose the edges of the graph that observe the constraints while leading to the lowest overall cost.

Dongmin et. al [13] propose an optical flow method for automatic cell tracking. The key algorithm of the method is to align an image to its neighbors in a large image collection consisting of a variety of scenes. Considering the method cannot solve the problems in all cases of cell movement, another optical flow method, SIFT (Scale Invariant Feature Transform) flow, is also presented. The experimental results show that both methods can track the cells accurately.

We wanted to have a cell tracking method that can use an operator at the cell line without an initialization phase or heavy human interaction. The method should be fast and robust enough and should not have any subjective human interaction. Based on our observation, we concluded that it must be possible to track the cells by identifying the same cell from one image to the next image. Therefore, we introduced a similarity-based approach [20] for cell tracking. This idea has been overtaken later on by Padfield et. al [8] and by Kan et. al [9]. Both approaches [8], [9] do not evaluate their approach with our approach.

Padfield et.al [8] use spatiotemporal volumes in order to keep track of moving cells over time. The methods uses a standard level set approach for segmentation to extract the G2 phase nuclei from the volume and a Euclidean distance metric for linking between the G2 phases across the other phases. The level-set approach allows to recognize cells at different granularity level but will not prevent the algorithm to detect changing morphology. The Euclidean distance is a simple similarity measure that does not take into account the specific characteristics of visual objects.

Given an estimate of the cell size only, the method of Kan et. al [9] is capable of ranking the trackers according to their performance on the given video without the need for ground truth. Using the cell size for tracking is a simple measure that is included in our application as well.

These two methods [8], [9] cannot track normal cells as well as mitotic cells as in our proposed approach. The Euclidean distance used as similarity measure in Kan et.al10 is not robust enough to handle changing cell morphology as it is in our case. We use a very flexible similarity measure [17] that considers the similarity of the corresponding pixels as well as the similarity to its surrounding pixels and therefore can handle morphological changes from one time-frame to the following time-frame. Cells that appear new in the microscopic image window can be tracked as well.

3. Material and Investigation of the Material

The live cell movies are comprised of a sequence of 113 images. These images are taken in a time interval of thirty minutes. Some of these cell sequences are shown in Fig. 1 but not in full time schedule. The whole experiment runs over two and a half days. The cells show no steady movement. They might suddenly jump from one direction to the opposite one. The cells might also turn around their own axis that will change their morphological appearance. They also might appear and disappear during the experiment as well as re-appear after having gone out of the focus of the microscope. It is also highly likely that the cells might touch or overlap. The tracking algorithm should run fast enough to produce the result in little computation time.
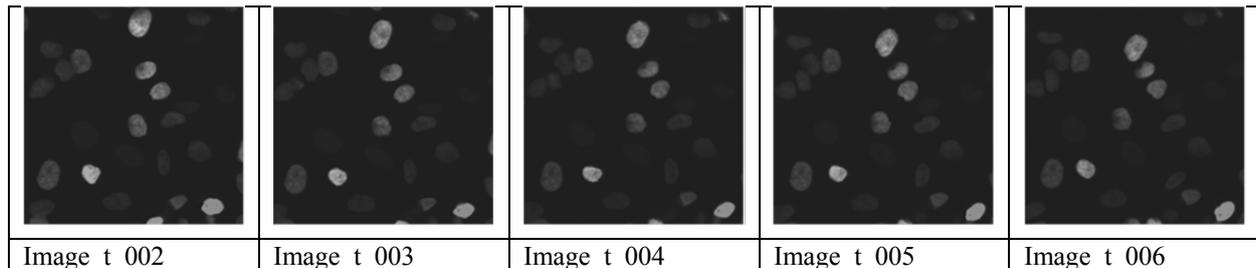

| Image_t_002 | Image_t_003 | Image_t_004 | Image_t_005 | Image_t_006 |

Fig. 1   Time Series of Image of a Live Cell Project

The following conditions have been decided for the algorithm: Only cells that are in focus should be tracked (background objects are not considered). Fragmented cells at the image borders are eliminated. Each detected cell gets an identity label and gets tracked. Disappearing cells are tracked until they disappear. Newly appearing cells are tracked upon their appearance and get a new identity label. Cells overlapping or touching each other are eliminated. They are considered as disappearing cells. Splitting cells are tracked after splitting and both get an identity label that refers to the mother cell. Cell fragments are eliminated by a heuristic less than 2xSigma of Cell Size. Note that we decided to exclude overlapping cells for the reason of the resulting higher computation time. If we want to consider this kind of cells, we can use our matching algorithm in [15] to identify the portion of touching cells belonging to the cell under consideration.

4. The Object-Tracking Algorithm

The unsteady movement of the cells and the movement of the cells around their own axis required special processing algorithms that are able to handle this. The successive estimation of the movement based on mean-shift or Kalman filters [15] would only make sense if we can expect the cells to have steady movement. Since it can happen that the cells jump from one side to another, we used a search window around each cell to check where the cell might be located. The size of the search window is estimated based on the maximal distance a cell can move between two time frames. The algorithm searches inside the window for similar cells. Since we cannot expect the cell to have the same shape after turning around its own axis and as also the texture inside the cell might change, we have chosen a similarity-based procedure for the detection of the cell inside the window. The overall procedure is shown in Fig. 2.
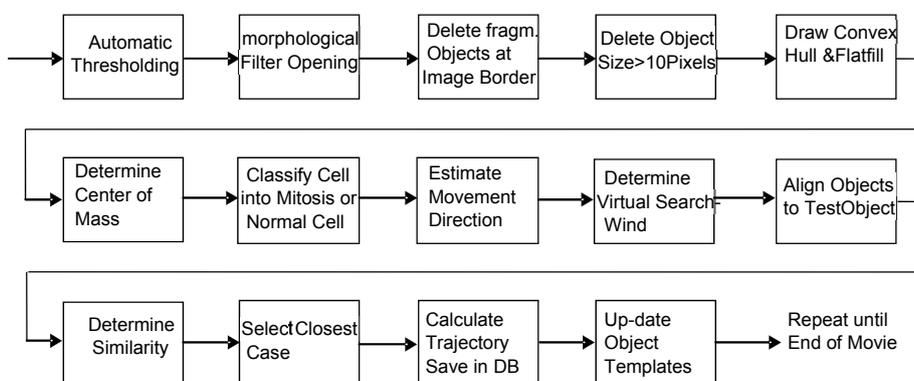

Fig. 2   Flowchart of the algorithm

4.1   Overall Algorithm

The image gets thresholded by Otsu`s well-known segmentation procedure. Afterwards the morphological filter opening by a 3x3 window is applied to close the contour and the inner holes. Fragmented cells at the image borders as well as small remaining objects of a size of ten pixels are deleted. The cells at the image borders are only tracked when they fully appear inside the image. Around the object is drawn the convex hull and remaining holes or open areas inside the cell area are closed by the operation flat-fill. The resulting images after these operations are shown in Fig. 3 for the six time frames. This resulting area is taken as the cell area and the area with its grey levels is temporarily stored as template in the data base.

   Then the center of gravity of the object is determined and the search window is tentatively spanned around the object. A raw estimation of the cell's movement direction is calculated by the mean-shift filter over 3 frames. In the resulting direction is started the search for similar cells. Cells fragmented by the window border are not considered for further calculation. Each cell inside the window is compared by the similarity measure to the respective template of the cell under consideration. Before the similarity is determined the cells are aligned, so that they have the same orientation. The cell having the highest similarity score to the template is labeled as the same cell moved to the position x,y in the image t+1. The template is updated with the detected cell area for the comparison in the next time frame. The position of the detected cell is stored into the data base under the label of the template. Mitotic cells are detected by classifying each cell based on the texture descriptor given in [16] and the decision tree classifier.

   Let CB_tB be the cell at time-point t and $CB_{t+1}B$ the same cell at time point t+1. Then the rule for labeling a cell as "disappeared" is: IF CB_tB has no $CB_{t+1}B$ THEN Disappearing_Cell.
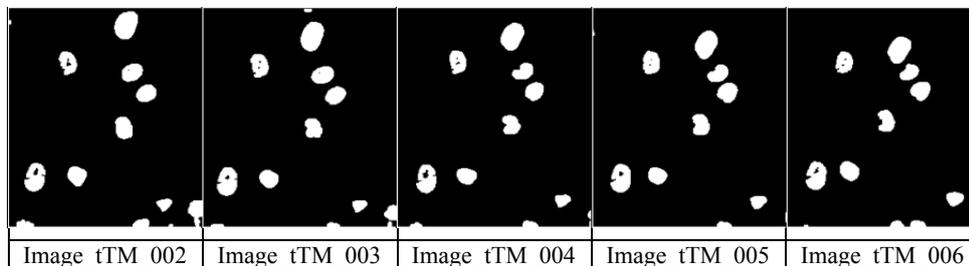

| Image_tTM_002 | Image_tTM_003 | Image_tTM_004 | Image_tTM_005 | Image_tTM_006 |

Fig. 3   Image after Thresholding and Morphological Filtering

4.2   Similarity Determination between two Cells

The algorithm [17] computes the similarity between two image areas A and B these images are in our case the bounding box around each cell (see Fig. 4). According to the specified distance function, the proximity matrix is calculated, for one pixel at position r,s in image A, to the pixel at the same position in image B and to the surrounding pixels within a predefined window. Then, the minimum distance between the compared pixels is computed. The same process is done for the pixel at position r,s in image B. Afterwards, the average of the two minimal values is calculated. This process is repeated until all the pixels of both images have been processed. The final dissimilarity for the whole image is calculated from the average minimal pixel distance. The use of an appropriate window size should make this measure invariant to scaling, rotation and translation.
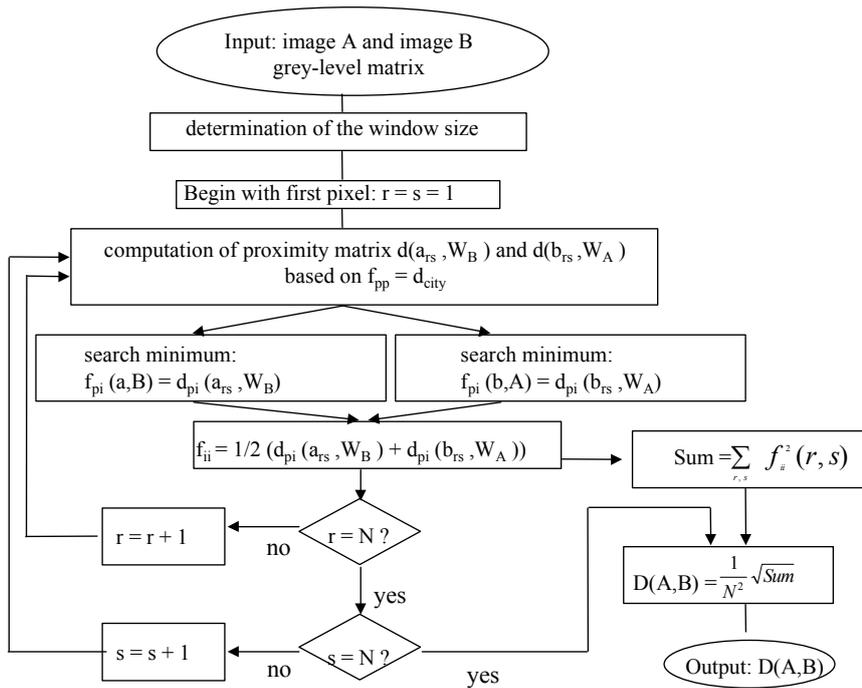
Fig. 4   Flowchart of the algorithm for computing the similarity measure.

The resulting similarity value based on this similarity measure for the pair wise comparison of cell_ 1 to cell_1, cell_1 to cell_2, and cell_1 to cell_3 in the preceding time frame is given in Fig. 5. The lowest similarity value is obtained when comparing cell_1 to cell_1 in the proceeding time frame.
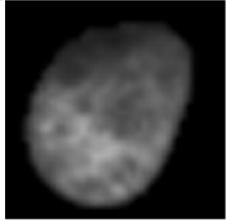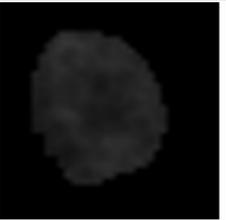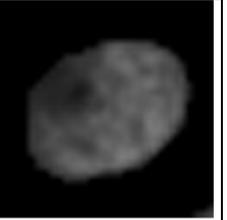

Cell_1 Image002 Template	Cell_1 Image003	Cell_2_Image003	Cell_3_Image003
Similarity value	**0.0554918423**	0.1190173551	0.0833996087

Fig. 5   Cells used for comparison

5. Results of Tracking Algorithm

Results in Fig. 6 a-e show the tracking path of six different cells. We compared the manually determined path of a cell from an operator of the High-Content Screening Process-line with the automatic determined path by the tool *IBaI-Track* for 10 videos with 117 frames each. If both methods gave the same path we evaluated it as positive otherwise as negative. We observed a correspondence between these two descriptions of 98.2 %. The computation time for a sequence of 117 images of 674x516 pixels each and on average 80 cells per image is 11 minutes 42 seconds on PC with 1.8 GHz.
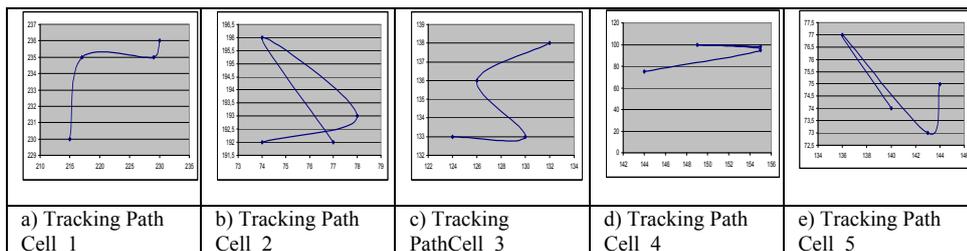

a) Tracking Path Cell_1	b) Tracking Path Cell_2	c) Tracking PathCell_3	d) Tracking Path Cell_4	e) Tracking Path Cell_5

Fig. 6   Cell tracking path

6. Information Extracted from the Path

The output of the cell-tracking algorithm is a tuple of coordinates for each cell that describes that path of the cell (see Fig. 6a-6e and Fig. 7). Biologists want to study the kinetics of the cells. Therefore, we have to extract features from this path that describe the motility and velocity of a cell. Table 1 shows features that are used to describe the path of a cell [18]. These features (see Table 2) are provided in a table to the biologist for further study. Please note, one image contains many cells. As result, we obtain a bunch of numerical values for one image that is hard to overlook for a human. More high-level descriptions are necessary that summarize these features and their feature values in a more compact information about the kinetics of the cells in one image. Recently, biologist use statistical tools to study these features. More complex image mining methods such as decision tree induction and conceptual clustering [19] can be of help in order to bring out the higher-level descriptions.
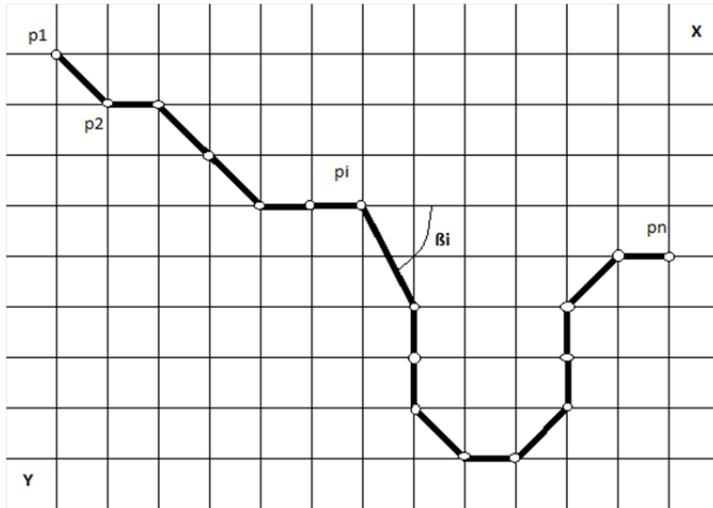
Fig. 7   Path of a Cell marked with coordinate point

Table 1   Measures for motility and velocity of a cell

Measure	Definition
Total distance traveled	$d_{\text{tot}} = \sum_{i=1}^{N-1} d(\mathbf{p}_i, \mathbf{p}_{i+1})$
Net distance traveled	$d_{\text{net}} = d(\mathbf{p}_1, \mathbf{p}_N)$
Maximum distance traveled	$d_{\text{max}} = \max_i d(\mathbf{p}_1, \mathbf{p}_i)$
Total trajectory time	$t_{\text{tot}} = (N-1)\Delta t$
Confinement ratio	$r_{\text{con}} = d_{\text{net}}/d_{\text{tot}}$
Instantaneous angle	$\alpha_i = \arctan(y_{i+1} - y_i)/(x_{i+1} - x_i)$
Directional change	$\gamma_i = \alpha_i - \alpha_{i-1}$
Instantaneous speed	$v_i = d(\mathbf{p}_i, \mathbf{p}_{i+1})/\Delta t$
Mean curvilinear speed	$\bar{v} = \frac{1}{N-1}\sum_{i=1}^{N-1} v_i$
Mean straight-line speed	$v_{\text{lin}} = d_{\text{net}}/t_{\text{tot}}$
Linearity of forward progression	$r_{\text{lin}} = v_{\text{lin}}/\bar{v}$
Mean squared displacement	$\text{MSD}(n) = \frac{1}{N-n}\sum_{i=1}^{N-n} d^2(\mathbf{p}_i, \mathbf{p}_{i+n})$

Table 2   Output of the Cell tracking Tool

Cell Number	Total Distance traveled	Maximum distance traveled	…	…	…	…	Mean squared displacement
Cell_1	X11	X12	…	…	…	…	X112
…	…	…	…	…	…	…	
Cell_n	Xn1	Xn2	…	…	…	…	Xn12

7. Conclusions

We have presented our new cell-tracking algorithm for tracking cells in dynamic drug discovery or computational biology experiments. The algorithm uses a window around the position of the cell in the preceding time frame for searching for the next position of the cell. This is necessary since the movement of the cells cannot be steady. The search inside the window is started in the expected direction of the cell. To calculate this direction, we use a mean-shift filter over 3 time frames. The detection of the cell is done based on a similarity determination of the grey-level profile between the cells in the preceding time-frame and the following time-frame. The cell giving the highest similarity value is selected as the same cell in the following time-frame. The template is updated with the new cell. A similarity measure that can handle small changes in the morphology of the cells is used. This similarity measure is robust enough to detect the cell with high accuracy. The tracking algorithm can track normal cells and mitotic cells by our formerly developed classifier. Eleven features describe the path of a cell. The resulting data are stored in a data file and can be used for further image-mining analysis. We propose that biologists and drug discovery experts think about more high-level terms to describe the events they want to discover since the large bunch of numerical values created during the tracking process is hard to overlook by a human. For that can be used proper image mining algorithm such as decision tree induction and conceptual clustering.

References

[1] O. Debeir, P.V. Ham, R. Kiss, C. Decaestecker, Tracking of migrating cells under phase-contrast video microscopy with combined mean –shift processes, IEEE Trans. Med. Im. 2005 (24), 697-711

[2] D. Padfield, J. Rittscher, T. Sebastian, N. Thomas, B. Roysam, Spatiotemporal cell segmentation and tracking in automated screening, In: Proc. of the IEEE Intern. Symp. Biomed. Imaging, 2008, pp 376-379

[3] K. Li, E.D. Miller, M. Chen, T. Kanade, L.E. Weiss, Ph.G. Campbell, Cell population tracking and lineage construction with spatiotemporal context, Medical Image Analysis, 12 (2008), p. 546-566

[4] K. Li, E.D. Miller, M. Chen, T. Kanade, L.E. Weiss, Ph.G. Campbell, Cell population tracking and lineage construction with spatiotemporal context, Medical Image Analysis, 12 (2008), p. 546-566

[5] Sacan, H. Ferhatosmanoglu, H. Coskun, CellTrack: an open-source software for cell tracking and motility analysis, BIOINFORMATICS APPLICATIONS NOTE Vol. 24 no. 14, 2008, pages 1647–1649

[6] M. Wang, X. Zhou, F. Li, J. Huckins, R. W. King, St. T.C. Wong, Novel cell segmentation and online SVM for cell cycle phase identification in automated microscopy, BIOINFORMATICS, Vol. 24 no. 1, 2008, pages 94–101

[7] Cohen, A.R. Bjornsson, C. Ying Chen, Banker, G., Ladi, E., Robey, E., Temple, S., Roysam, B. Automatic summarization of changes in image sequences using algorithmic information theory, 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2008, 859-862

[8] D. Padfield, J. Rittscher, N. Thomas, and B. Roysam, Spatio-Temporal Cell Cycle Phase Analysis Using Level Sets and Fast Matching Methods, In: Proc. Of the IEEE Intern. Symp. Biomed. Imaging, 2008, pp 376-379

[9] Andrey Kan, John Markhamb, Rajib Chakravortyb, , James Baileya, , Christopher Leckiea, Ranking cell tracking systems without manual validation, Pattern Recognition Letters, Volume 53, 1 February 2015, p. 38–43

[10] Martin Maška, Ondřej Daněk, Saray Garasa, Ana Rouzaut, Arrate Muñoz-Barrutia, and Carlos Ortiz-de-Solorzano, Segmentation and Shape Tracking of Whole Fluorescent Cells Based on the Chan–Vese Model, IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 32, NO. 6, JUNE 2013, 995-1005

[11] Magnusson K.E.G., Jalden J. ,Gilbert P.M., Blau, H.M., Global Linking of Cell Tracks Using the Viterbi Algorithm, Medical Imaging, IEEE Transactions on (Volume:34 , Issue: 4), November 2014, p. 911 – 929

[12] Dirk Padfield, Jens Rittscher, Badrinath Roysam, Coupled minimum-cost flow cell tracking for high-throughput quantitative analysis, Medical Image Analysis, Aug. 2011, Volume 15, Issue 4, p. 650–668

[13] Dongmin Guo, Anne L. Van de Ven, and Xiaobo Zhou, Tracking and Measurement of the Motion of Blood Cells Using Optical Flow Methods, IEEE J Biomed Health Inform. 2014 May; 18(3): 991–998.

[14] P. Perner, H. Perner, and S. Jänichen, Recognition of Airborne Fungi Spores in Digital Microscopic Images, Journal Artificial Intelligence in Medicine, Volume 36, Issue 2 , February 2006, p. 137-157

[15] D. Comaniciu, V. Ramesh, and P. Meer. Real-time tracking of non-rigid objects using mean shift. In Proc. IEEE Conf. Comp. Vision and Patt. Recog., volume 2, pages 142–149, June 2000

[16] Anja Attig and Petra Perner, A Comparison between Haralick's Texture Descriptor and the Texture Descriptor based on Random Sets for Biological Images, In: Petra Perner (Ed.) Machine Learning and Data Mining in Pattern Recognition. lnai 6871, pp.524-538, Springer Verlag 2011

[17] Zamperoni P., Statovoitov V., How dissimilar are two gray-scale images, In: Prod. Of the 17th DAGM Symposium, Springer Verlag Berlin, 1995, pp. 448-455

[18] Beltman JB, Mar´ee AFM, de Boer RJ (2009). Analysing immune cell migration. Nature Reviews Immunology 9, 789–798.

[19] Perner P., Data Mining on Multimedia Data, Lecture Notes in Artificial Intelligence, Vol. 2558. Springer Verlag, Berlin Heidelberg New York (2002), ISBN: 3-540-00317-7

[20] H. Perner and P. Perner, Similarity-Based Motion Tracking of Cells in Microscopic Images, IBaI Report 2006, IB2006-11

[21] P. Perner, Progress in Cell Image Analysis and Life Cell Tracking, QuantPro-BMWI Presentations and Reports May/2006, September/2007, March/2008 grant no. 0313831B