

Determination of Mitotic Delays in 3D Fluorescence Microscopy Images of Human Cells using an Error-Correcting Finite State Machine

Nathalie Harder¹, Felipe Mora-Bermúdez², William J. Godinez¹,
Jan Ellenberg², Roland Eils¹ and Karl Rohr¹

¹University of Heidelberg, IPMB, and DKFZ Heidelberg, Dept. Bioinformatics and
Functional Genomics, Im Neuenheimer Feld 364, D-69120 Heidelberg,

²European Molecular Biology Laboratory (EMBL), Gene Expression and Cell
Biology/Biophysics Programmes, Meyerhofstrasse 1, D-69117 Heidelberg
Email: n.harder@dkfz-heidelberg.de

Abstract. In high-throughput cell phenotype screens large amounts of image data are acquired. The evaluation of these microscopy images requires automated image analysis methods. Here we introduce a computational scheme to process 3D multi-cell image sequences as they are produced in large-scale RNAi experiments. We describe an approach to automatically segment, track, and classify cell nuclei into seven different mitotic phases. In particular, we present an algorithm based on a finite state machine to check the consistency of the resulting sequence of mitotic phases and to correct classification errors. Our approach enables automated determination of the duration of the single phases and thus the identification of cell cultures with delayed mitotic progression.

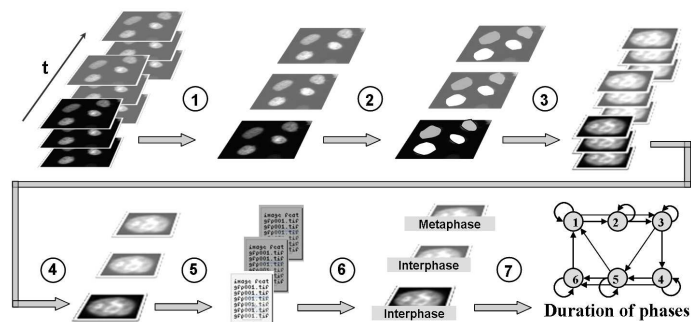
1 Introduction

RNA interference (RNAi) is an effective tool for identifying the biological function of genes. With this method genes are systematically silenced and the resulting morphological changes are analyzed. However, such large-scale screens provide large amounts of data which require tools for automated image analysis.

Our work is carried out within the project MitoCheck, which has the goal to explore the processes of cell division (mitosis) in human cells at a molecular level. In this project RNAi secondary screens are performed and fluorescence microscopy image sequences of the treated cell cultures are acquired to study the effects of the silenced genes on mitosis. This contribution is concerned with the automated evaluation of an assay for studying delays in mitotic phases, which are caused by gene silencing. Thus, the duration of the different phases of cell division has to be measured for the treated cells and compared with the normal cells from control experiments. Therefore, cells have to be observed throughout their life cycle and for each time point the respective phase has to be determined.

Automated analysis of high-throughput cell phenotype screens plays an increasingly important role. Approaches to analyse single-frame multi-cell 2D images from large-scale experiments have been described, e.g., in [1]. Recently,

Fig. 1. Image analysis workflow: (1) Maximum intensity projection of multi-cell 3D images, (2) Segmentation and tracking in 2D, (3) Extraction of 3D ROIs that include single cells, (4) Selection of most informative slices, (5) Feature extraction, (6) Classification, (7) Consistency check, error correction, and determination of phase durations



work has been done on automatically processing multi-cell 2D image sequences for cell cycle analysis. With these approaches, cells are segmented, tracked, and classified given phase-contrast [2] or fluorescence [3] microscopy images. In [2, 3] cells or cell nuclei are classified into a maximum of four phases. In [4] 3D image sequences are processed and cell nuclei are classified into seven cell cycle phases but no consistency check of the resulting phase sequences has been performed.

Our approach allows to analyze 3D multi-cell image sequences from a confocal fluorescence microscope. We classify cell nuclei into seven mitotic phases. To this end, we have developed a workflow that comprises segmentation, tracking of splitting nuclei, extraction of static and dynamic features, and classification. In particular, we present a scheme to check the consistency of the resulting sequences of mitotic phases of cells based on biological constraints. This scheme is based on a finite state machine which enables to automatically resolve errors using the confusion probabilities of the classifier. As a result, we can determine the lengths of the seven mitotic phases of cell nuclei automatically.

2 Methods

Image analysis workflow To analyze the mitotic phases high-resolution confocal fluorescence microscopy 3D images of the DNA are acquired which consist of three confocal planes (slices). Because of technical reasons in the image acquisition process the number of slices is restricted to three. During mitosis the cell changes its shape (gets more rounded) and therefore in different phases the DNA is visible in different slices. Using multi-cell images that contain cells in different mitotic phases, it is impossible to define one slice per time step that well represents the DNA of all cells. Therefore, we have developed the workflow shown in Fig. 1. First, we apply a maximum intensity projection (MIP) for each time step, resulting in 2D images (Fig. 2 (left)). Based on these MIP images we

Fig. 2. (left) Original 3D image (Maximum intensity projection), (right) 1-4: Example for the tracking of a mitotic nucleus in four consecutive time-steps, 4a: Result without mitosis detection, 4b: Result after mitosis detection and track merging

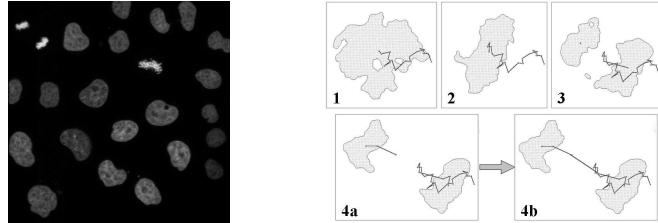


Fig. 3. Example images for the seven different mitotic phases (from left to right): inter-, pro-, prometa-, meta-, ana- 1, ana- 2, and telophase

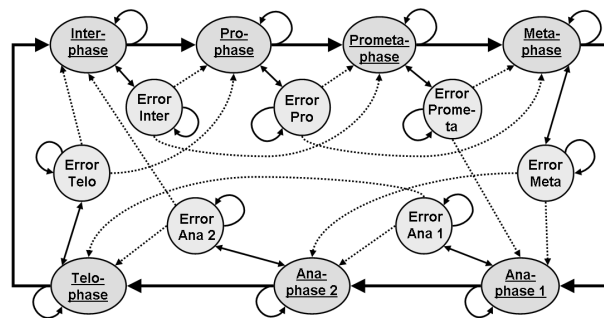


perform segmentation and tracking to determine the correspondences in subsequent frames. We now go back to the 3D images and define 3D ROIs for each cell based on the segmentation and tracking result. For each 3D ROI we choose the most informative slice and extract static and dynamic features. Then we apply a classifier which results in a sequence of mitotic phases for each cell trajectory. Finally, the resulting phase sequences are parsed with a finite state machine to check their consistency, resolve inconsistencies, and determine the phase lengths.

Segmentation and tracking of mitotic cell nuclei For segmentation we apply region-adaptive thresholding which proved to be fast and robust in our application. This scheme computes local intensity thresholds in overlapping image regions using histogram-based threshold selection. To analyze the mitotic behavior of single cells, a tracking scheme is required that determines the temporal connections for splitting objects. We have developed the following two-step scheme: First, initial, non-splitting trajectories are established, and second, mitotic events are detected and the related trajectories are merged resulting in tree-structured trajectories (Fig. 2 (right)). For more details see [4].

Feature extraction and classification To compute image features we select for each cell nucleus its individual most informative slice. Our experiments showed that the maximum total intensity performs very well as selection criterion. Within the most informative slice we compute static and dynamic image features. The static features comprise object- and edge-related features, texture features, grey scale invariants, and Zernike moments. As dynamic features we compute the difference of object size, intensity mean and standard deviation, and circularity for each nucleus to its predecessor and successor. We apply a support vector machine (SVM) classifier with a radial basis function (RBF) ker-

Fig. 4. Finite state machine to check the consistency of the computed phase sequences



nel to classify the nuclei into the classes interphase, prophase, prometaphase, metaphase, anaphase 1, anaphase 2, and Telophase (Fig. 3).

Consistency check and error correction In order to determine the phase lengths automatically it is necessary that the computed phase sequences are consistent (which is not always the case due to classification errors). Therefore, we have developed a finite state machine (FSM) that accepts only biologically possible phase sequences. Each phase is represented by one state of the FSM. The possible phase transitions are represented by the state relations. Fig. 4 shows a sketch of the FSM; for clarity only the most important relations are displayed. The FSM contains error states to resolve inconsistencies. An error state is quit if the following phase is a valid phase. Then the most likely phase for the error state is determined based on the confusion probabilities of the classifier that have been established using cross-validation on the training set. In one run of the FSM inconsistencies of length 1 are completely corrected. Errors of length larger than 1 can be successively resolved in multiple runs of the FSM. In addition, the FSM determines the phase durations.

3 Results

Our experiments are based on four multi-cell 3D image sequences each consisting of 124 time steps. 29 cell nuclei have been segmented and tracked over 124 time steps resulting in 4225 single cell 3D image stacks (note that the cells proliferate). Some few trajectories have been manually corrected since our tracking scheme currently detects 80% of the occurring mitoses. We have performed a kind of two-fold cross-validation on the 29 available trajectories where in each loop 18 tracks have been used for training and 11 tracks for testing. In each cross-validation loop first the confusion probabilities were determined using a five-fold cross-validation on the training set. Then, the classifier was trained with the whole training set and tested with the test set. We obtain an overall classification

Table 1. Sample numbers and classification accuracies for each class (22 tracks tested)

	Inter	Pro	Prometa	Meta	Ana1	Ana2	Telo
No. of samples	2986	51	36	76	14	50	152
Class. accuracy	99.0%	84.3%	94.4%	79.0%	57.1%	86.0%	69.7%

accuracy of 96.6% (the sample numbers and accuracies per class for both cross-validation loops are given in Tab. 1). The resulting phase sequences for the 22 classified trajectories were processed with the finite state machine using the confusion matrices determined on the training sets to resolve consistency errors. Finally, the corrected phase sequences have been compared with the manually assigned correct phases (ground truth). It turned out that all inconsistencies of length 1 have been resolved in the first run of the FSM. Applying the FSM a second time on the corrected sequences resolved also all inconsistencies of length 2. Longer inconsistencies have partly been resolved. Note that this correction step is essential for determining the phase lengths automatically.

4 Discussion

We have presented an approach for automated analysis of the duration of mitotic phases in 3D confocal microscopy image sequences. Our approach segments and tracks splitting nuclei and thus determines cell pedigrees. By using static and dynamic features our scheme classifies the cells into seven mitotic phases. The consistency of the computed phase sequences is checked and inconsistencies are resolved using a priori knowledge. Finally, the phase durations are determined.

In future work, we plan to improve the performance of the tracking and classification schemes by including additional image features. We will also apply our approach to a larger number of image sequences and evaluate its performance.

Acknowledgment

This work has been supported by the EU project MitoCheck.

References

1. Perlman ZE, Slack MD, Feng Y, et al. Multidimensional drug profiling by automated microscopy. *Science* 2004;306:1194–1198.
2. Yang F, Mackey MA, Ianzini F, et al. Cell segmentation, tracking, and mitosis detection using temporal context. *LNCS* 2005;3749:302–309.
3. Padfield DR, Rittscher J, Sebastian T, et al. Spatio-temporal cell cycle analysis using 3D level set segmentation of unstained nuclei in line scan confocal fluorescence images. *Procs ISBI* 2006; 1036–1039.
4. Harder N, Bermúdez F, Godinez WJ, et al. Automated analysis of the mitotic phases of human cells in 3D fluorescence microscopy image sequences. *LNCS* 2006;4190:840–848.