HKI-GE

Authors: Ivan Belyaev, Jan-Philipp Praetorius, Anna Medyukhina, Marc Thilo Figge

Email: thilo.figge@leibniz-hki.de

Platform: Linux (tested on Ubuntu 18.04)

Prerequisites: JSON for modern C++>=2.1.1, Boost>=1.65, OpenMP==5.0.1, OpenCV>= 3.2, LEMON Graph library>= 0.7

HKI-GE: SUMMARY

The algorithm for migration and interaction tracking version 3 (AMIT-v3) [1] is designed for segmentation and tracking of host-pathogen confrontation assays for their quantitative characterization. The number of migrating cells within the field of view should not be too high (i.e., for a typical number of up to roughly 60 cells per field of view of size 0.42 × 0.42 mm). The algorithm comprises two independent components: segmentation and tracking. The segmentation part allows to detect regions of interest (ROI) (e.g., isolated cells and cell clusters) in each frame using classical image processing techniques. The tracking part is organized in three steps. In the first step, clusters are detected by monitoring the events of cell-cell fusion and cluster fission into single cells both forward and backward in time. In the second step, all detected clusters are split using a multilayer cluster splitting procedure. Finally, in the third step, the acquired cells are tracked base on nearest neighbor association (NNA).

HKI-GE: PREPROCESSING

This step involves image contrast enhancement using top-hat and bottom-hat transformations with filter diameter *sd1_kernelSize* and *sd2_kernelSize*, respectively, and suppression of background variation by an adaptive Wiener filter.

HKI-GE: SEGMENTATION

This algorithm was developed for single-channel grayscale images; it processes each frame of a video individually and is based on classical image processing techniques [2]:

(i) detection of regions with relatively high intensity variation by utilizing a standard deviation filter followed by global thresholding with a fixed threshold *threshold_binary*;

(ii) morphological filtering to suppress objects on the frame borders, to remove objects smaller than *min_region_size* pixels, and to smooth object contours by erosion with filter diameter of *erode_kernelSize* pixels followed by small artefacts removal operation.

HKI-GE: TRACKING

The tracking routine uses binarized images. Its first step is the cluster identification based on a fusion-and-fission-algorithm for ROIs that is organized as follows: All frames of a video are analyzed one after the other in subsequent pairs (1st frame, 2nd frame), (2nd frame, 3rd frame), etc. Within each pair of frames, the predecessors-successors relationship for each segmented ROI is analyzed and recorded. We consider a ROI as a cluster, if (i) it overlaps with at least two ROIs from the previous frame and (ii) its size of intersection with each of them exceeds a certain threshold (*overlap_threshold*, decimal fraction of overlap area). The choice of this threshold value depends on the time resolution of time-lapse imaging as well as the activity of the cells. This allows preventing false overlap detections for data with low time resolution. Each fusion or fission event leads to an ID change for the current ROI. The result is a set of tracklets, where each object is assigned a type (i.e., single cell or cell cluster). Based on this information a map of detected clusters is generated for each frame. A second set of cluster maps is produced by the same sequence of operations, but for the reversed images order. At the end, the final cluster maps for each image are generated by combining the clusters identified in the forward and backward maps.
Then every cluster is processed independently within an area restricted by the corresponding bounding box using the cut-out of the original grayscale image. The three steps of the hierarchical cluster splitting are detailed next.

Single-cell segmentation. The first step of hierarchical cluster splitting deals with the analysis of cases where boundaries between cells within a cluster are easily distinguishable. On the cut-out of the respective grayscale image, contrast enhancement will be performed to emphasize the boundaries of every cell. To restrict the noise level, Gaussian filtering is applied, followed by Canny edge detection of the cell boundaries with threshold *canny_threshold* (in units of intensity) and morphological closing and opening operations to fill holes and to remove small artifacts. To ensure that the detected objects overlap with the original binary cluster detection region, a binary overlap is performed between these two images. At this point, the number of underlying objects resulting from the single-cell segmentation are compared with the number of known cells for the corresponding cluster. In case the two numbers are equal, watershed segmentation will be executed with seed regions as obtained from the single-cell segmentation.

Distance transformation. In cases where the membrane of the respective cells within a cluster is blurred or not detectable, single-cell segmentation alone is not enough to obtain the correct number of seed

regions. Sometimes only a tiny connection remains between two or more cells within a cluster. For these and all other remaining connections, where the number of seed regions does not match the known number of single cells within a cluster, a distance transformation is applied. This procedure requires a binary image as input and computes for each foreground pixel the distance to the next background pixel with an appropriate distance metric. In order to substantiate this step qualitatively, the distance transformation is performed on both the original binary image and the resulting mask from the previous step. The results of the two parallel processes usually differ only very slightly and reach the required number of markers, so that this step is aborted, and a watershed segmentation is performed. In AMIT-v3, we used the Euclidean distance metric and then normalized all distances relative to the maximum distance. Next, we thresholded all distances by setting the associated pixels to background pixels, if the underlying distance from the foreground to the next background pixel is shorter than a certain threshold. Here we chose a threshold of 40 percent of the normalized distances. For some clusters, especially for those that are circular-shaped, this step must be repeated until the number of seed regions matches the number of known cells within a cluster, or until the number of repetitions reaches a predefined maximal number. In most cases, the required number of seed regions for a cluster is achieved by the distance transformation and the splitting process can be completed.

Random seed region generation. In some cases, the required number of seed regions may not be obtained and require to randomly generate seed regions according to the known number of cells within the cluster under consideration. For up to five known cells in a cluster, the random seed regions are extracted from the top/bottom/right/left-most point and the center point inside the cluster. If there are more cells in the cluster, the additionally required seed regions are extracted completely randomly from the cluster. However, this procedure is just an exception handling if the conventional methods do not work, and this is usually only rarely required. Once all seed regions have been initialized, we perform cluster splitting through the watershed segmentation. The obtained seed regions are used as markers for a marker-based watershed segmentation, which is performed on the original binary mask of the current cut-out image. This results in splitting the cut-out mask into as many regions as there are cells in the cluster, as is known from the cluster detection algorithm. Finally, the split cut-out image is placed at its respective position within the original full-size image.

HKI-GE: POST-PROCESSING

No post-processing step has been taken after tracking.

REFERENCES

1. https://github.com/applied-systems-biology/amit
2. Gonzalez RC, Woods RE. Digital Image Processing (2007).