Large scale tracking of stem cells using sparse coding and coupled graphs

Vestergaard, Jacob Schack; Dahl, Anders Bjorholm; Holm, Peter; Larsen, Rasmus

Publication date:
2012

Document Version
Publisher’s PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Abstract

Stem cell tracking is an inherently large-scale problem. The challenge is to identify and track hundreds or thousands of cells over a time period of several weeks. This requires robust methods that can leverage the knowledge of specialists on the field. The tracking pipeline presented here consists of a dictionary learning method for segmentation of phase contrast microscopy images. Linking of the cells between two images is solved by a graph formulation of the tracking problem.

This poster is prepared as part of the PhD thesis entitled “A Grand Challenge: Large Scale Event Recognition and Tracking.”

The phase microscopy images are kindly provided by the Department of Basic Animal and Veterinary Sciences, Faculty of Life Sciences, Copenhagen University.

Dictionary learning

Learning a discriminative dictionary image patches is a flexible and intuitive method for supervised segmentation. The idea is to build a basis for classification from the type of data at hand, rather than using a fixed basis, such as wavelets, convoluted, or similar.

The dictionary atoms can be configured to contain any type of features. E.g., for multispectral data — or data in multiple modalities — the dictionary atoms can contain all channels. These will simply be part of the vectorized atom.

Learning a discriminative dictionary from an image:
1. Label an image with a number of classes.
2. For each pixel, select an image patch.
3. Cluster these image patches in an intensity dictionary and an associated label dictionary with probabilities of each class.
4. Increase the discriminative power of the dictionary, by further separating cluster centers.

To segment an image using the dictionary is simple: Each image patch in the image is “looked up” in the intensity dictionary using a nearest neighbor search and assigned the corresponding probabilities from the label dictionary. For overlapping image patches, the probabilities are accumulated.

The result of the dictionary segmentation is a probability map with a pixel probability for each class.

High dynamic range (HDR) phase contrast microscopy

Acquisition of the best possible image is ensured using high dynamic range (HDR) correction of the images. The phase contrast microscope is setup to acquire images in 4 by 4 grid. In each of these observation points, four exposures are collected ranging from over exposed to under exposed. The over exposed images possess a large amount of saturated pixels, but captures the nuances in the center of the soma. The under exposed images neglects these nuances, but decreases the extent of the halo surrounding the soma.

The response function g(\(\eta\)) is modeled as a smooth monotonic function of the irrelevances and is recovered by minimization of the quadratic objective function

\[ \sum_{i,j} \alpha_{ij} (g(E_{ij}) - g(L_{ij}))^2 + \lambda \sum_{i,j} f_{ij}^2 \]

where \(E_{ij}\) is the exposure time of the \(i\)th pixel, \(L_{ij}\) the unknown irrelevances in the \(i\)th pixel and \(f_{ij}\) the pixel value. Smoothness is enforced by regularization of the second order derivative in the pixel intensity range.

Having recovered g(\(\eta\)) the four images are combined into a single image of irrelevances:

\[ L_{ij} = \sum_{k} \alpha_k E_{ij} - \log(\alpha_k) \sum_{k} \alpha_k f_{ij} \]

Where w(\(\eta\)) is an introduced weighting function ignoring saturated pixels. By this transformation the dynamic range is effectively increased.

Coupled graphs for topology changes

Tracking cells through the entire time series is solved using a graph formulation of the correspondence problem. The minimum cost solution to the graph can be obtained using linear programming.

BLDDs detected in two adjoining images are the nodes connected in the graph. The edge capacities are set such that the flow conservation constraint (flow in \(\rightarrow\) flow out) is honored and a minimum flow constraint is enforced.

The graph is set to accommodate for cell behavior and topology changes. A cell is allowed to:
- Move
- Appear
- Disappear
- Split to two cells
- Merge from two cells

The split behavior is implemented by coupling two edges on the graph: a pseudo edge from the appear node to the split node and an edge from the cell to the split node. Similarly for the merge behavior. This coupling is done to ensure that the flow conservation constraint is honored.

Defining the edge costs

The edge cost is a scalar representing the cost of including an edge in the solution. The cost can therefore be defined from the similarity of the two objects. To represent the similarity, the feature vector difference is used. The features included are position in the x- and y-direction and area.

The edge cost is defined as the Mahalanobis distance from this difference vector to the mean of a population of typical charged.

\[ \cos(c) = \frac{(S_{ij} - \mu_{ij})^T \Sigma^{-1} (S_{ij} - \mu_{ij})}{\sqrt{(S_{ij} - \mu_{ij})^T \Sigma (S_{ij} - \mu_{ij})}} \]

where \(S_{ij} = E_{ij} - L_{ij} \)

The parameters \(\mu_{ij}\) and \(\Sigma\) are estimated from a simple nearest-neighbor tracking using the principle of Full Width at Half Maximum (FWHM) to extract the dominant distribution.

Future work

Near future work include collection of ground truth to quantitatively evaluate the cell linking and definition of interesting features to collect.

The motivation for the stem cell tracking is to extract measures from the stem cell cultivation, enabling the bio analysts to optimize the growth process. Measures could include cell-to-cell proximity, in order to determine the cause for unexpected growth patterns or malious cells.

A longer-term perspective is n-coloring of specific cell types at the end of the cultivation process, ideally, the collected cell lineage will 1) enable the analyst to back track the life of each cell of interest back in time and 2) contribute to statistics of the cell types of cells.

References


DTU Informatics
Department of Informatics and Mathematical Modelling