TRACKING BOID PARTICLES USING MOTILE

1 Introduction

Two time-lapse datasets of 501 time frames each (referred here, as *Sequence A* and *Sequence B*), showing appearing, disappearing and dividing particles were simulated, and made available here: `https://www.dropbox.com/work/cell_tracking/set_24_oct_16`. Position, ground truth velocity and cell types for these particles were provided as features. All particles belong to one of 16 cell types.

We investigate how well MOTILE [1] tracks these simulated particles by employing different subsets of particle features (see Table 1) and using some hand-crafted weights, to weigh individual, employed features. The tracked result is then compared with the ground truth tracks (available, since the particles were simulated) to obtain evaluation metrics, which were computed using the TRACCURACY [2] python package.

2 Candidate Graph, Costs and Constraints

Each particle at frame t was linked to its 5 nearest spatial neighbors at $t + 1$ to produce a candidate graph. Next, *hyper edges* were added: a hyper edge takes the form $((u,), (v_1, v_2))$ where u is the node id of the particle at frame t and v_1 and v_2 are the node ids of particles at frame $t + 1$, and were used to indicate potential divisions. We add a binary indicator variable on each edge (X_e), and for node (particle) appearance (X_a) and disappearance (X_d). The binary variables $X = \{X_e, X_a, X_d\}$ are optimized over, to minimize an objective C:

$$\hat{X} = \arg\min C = \arg\min C_e X_e + C_a X_a + C_d X_d, \tag{1}$$

where C_e is the cost of selecting an edge and is a linear sum of distances on various node (particle) features, C_a is the cost of starting a new track and C_d is the cost of ending an existing track. This cost C is minimized under biologically feasible constraints (for example, each particle at frame t can have at most one predecessor at frame $t - 1$ and at most two successors at the frame $t + 1$).

The cost of selecting an edge C_e can be expressed as a linear, weighted sum of distances d on various node (particle) features:

$$C_e = \sum_{i}^{K} \left(w_i d_i + b_i \right). \tag{2}$$

Here, K is the number of employed features, w_i and b_i are the weights used for the ith feature, and d_i is the distance calculated on the ith feature.

For both the experiments, we normalize all distances so that they have 0 mean and 1 standard deviation. For the first experiment, we use $1/0$ binary weights for all distances. For example, when considering position or cell-type, w is set to 1 and b is set to 0.

In order to compute distances on the cell type feature, the cell type is converted to a one-hot representation and then L_2 norm is computed on the one-hot representation.

3 Results

Table 1: Quantitative evaluation. For each dataset, we compare results of different combinations of features

Position	Velocity	Cell Type	TRA [3]↑	AOGM [3, 4]↓	FP$_e$[3] ↓	FN$_e$[3] ↓	WS$_e$[3] ↓	FP$_d$ ↓	FN$_d$ ↓	TP$_d$ ↑	MBC [5] ↑
					Sequence A						
✓	x	x	0.999	212.5	14	37	0	1	0	318	0.996
✓	✓	x	0.999	212.5	14	37	0	1	0	318	0.996
✓	✓	✓	0.999	212.5	14	37	0	1	0	318	0.996
					Sequence B						
✓	x	x	0.999	222.0	19	40	0	1	0	307	0.996
✓	✓	x	0.999	222.0	19	40	0	1	0	307	0.996
✓	✓	✓	0.999	222.0	19	40	0	1	0	307	0.996

4 Next Steps

As a next experiment, one could learn the weights using Structured SVM (as described here `https://funkelab.github.io/motile/learning.html`) instead of using hand crafted weights.

The current setup assumes that there are no false positive and false negative detections. In order to handle false positive detections, one could add a node selection cost, based on the confidence for that detection (output from a deep learning model). In order to handle false negative detections, one could introduce *skip edges*. This allows the optimization to select connections, which are not only between consecutive frames.

References

[1] Jan Funke, Talley Lambert, Caroline Malin-Mayor, Benjamin Gallusser, Florian Jug, and Ana Cristina Pascual Ramos, "Motile: Multi-object tracker using integer linear equations," 2024.

[2] Morgan Schwartz, Caroline Malin-Mayor, and et al, "Traccuracy," 2023.

[3] Martin Maška, Vladimír Ulman, Pablo Delgado-Rodriguez, Estibaliz Gómez-de Mariscal, Tereza Nečasová, Fidel A Guerrero Peña, Tsang Ing Ren, Elliot M Meyerowitz, Tim Scherr, Katharina Löffler, et al., "The cell tracking challenge: 10 years of objective benchmarking," *Nature Methods*, vol. 20, no. 7, pp. 1010–1020, 2023.

[4] Pavel Matula, Martin Maška, Dmitry V Sorokin, Petr Matula, Carlos Ortiz-de Solórzano, and Michal Kozubek, "Cell tracking accuracy measurement based on comparison of acyclic oriented graphs," *PloS one*, vol. 10, no. 12, pp. e0144959, 2015.

[5] Kristina Ulicna, Giulia Vallardi, Guillaume Charras, and Alan R Lowe, "Automated deep lineage tree analysis using a Bayesian single cell tracking approach," *Frontiers in Computer Science*, vol. 3, pp. 734559, 2021.