School of Computer and Communication Sciences

CVLAB


ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

K. Smith, A. Carleton, V. Lepetit



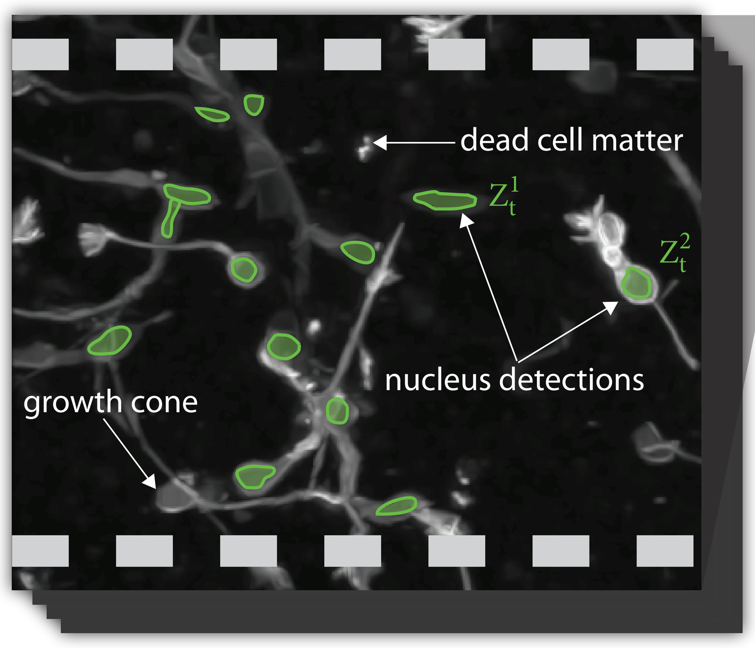
June 25, 2008

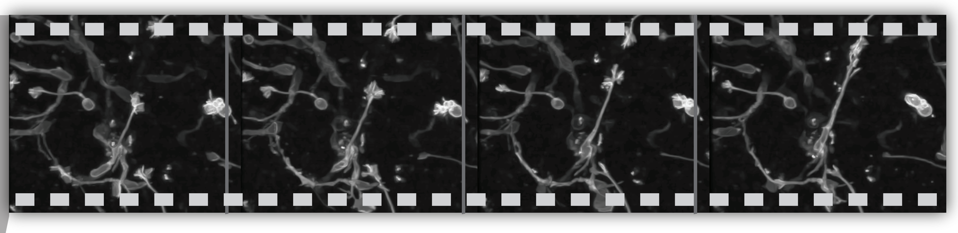

# General Constraints for Batch Multiple-Target Tracking Applied to Large-Scale Videomicroscopy

## Introduction

We present a principled probabilistic formalization of batch MTT introducing two general constraints to the tracking problem. The first constraint enforces a learned correlation between a target's appearance and its motion. The second constraint encourages proposed target paths to enter and exit near scene boundaries.

## Problem Formulation

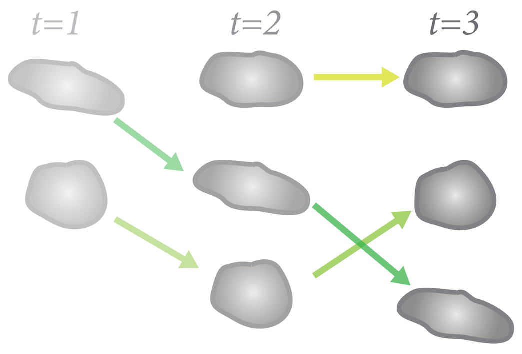


A set of nucleus detections $\mathcal{Z}_T$ is extracted from a time-series of image stacks.

Neuron nuclei are detected as dark, compact contours surrounded by a bright membrane.

Our goal is to link detections $\mathcal{Z}_T$ to form the most likely set of target paths $\mathcal{X}_T$: $\arg\max_{\mathcal{X}_T} p(\mathcal{X}_T|\mathcal{Z}_T)$

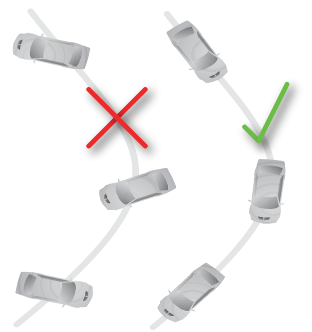



$$p(\mathcal{X}_T|\mathcal{Z}_T) \propto p(\mathbf{Z}_1|\mathbf{X}_1)\, p(\mathbf{X}_1) \prod_{t=2..T} p(\mathbf{Z}_t|\mathbf{X}_t)\, p(\mathbf{X}_t|\mathbf{X}_{t-1}) \quad (1)$$

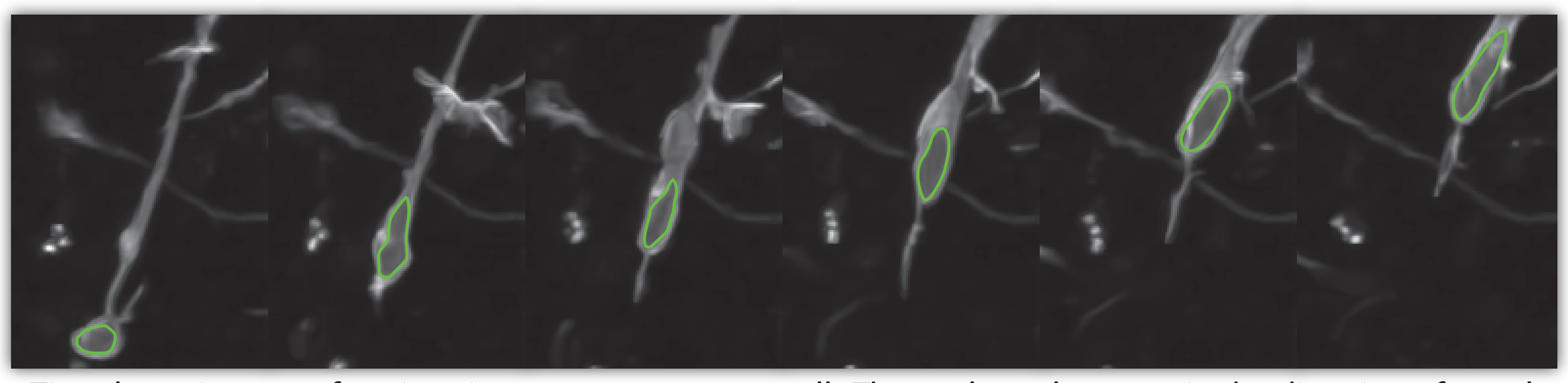
## The First Constraint

**Constraint 1**: The movement of a target and its appearance are not necessarily independent.

We relax the conventional assumption that motion is independent from appearance and learn a motion-appearance correlation model to assist tracking.



For some objects, motion and appearance are correlated.



Time-lapse images of a migrating neuron precursor cell. The nucleus elongates in the direction of travel.

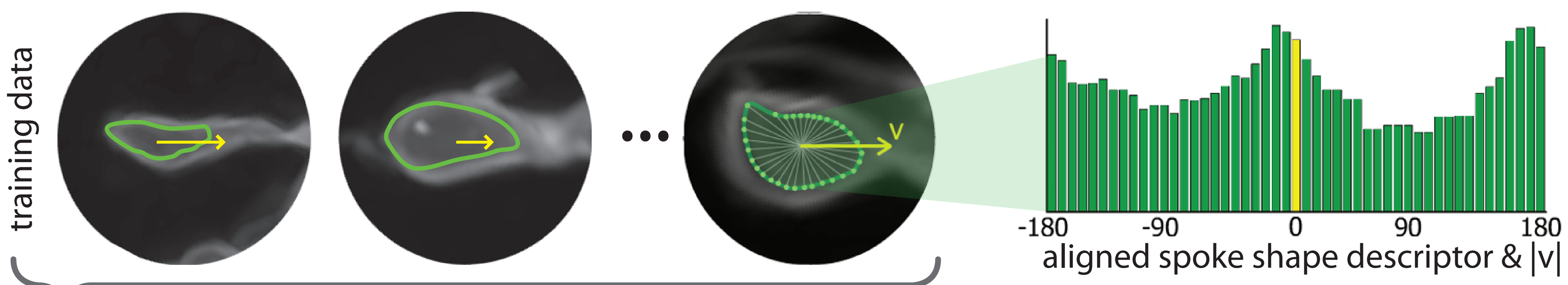
Constraint 1 gives rise to a new term in the observation model from $p(\mathcal{X}_T|\mathcal{Z}_T)$

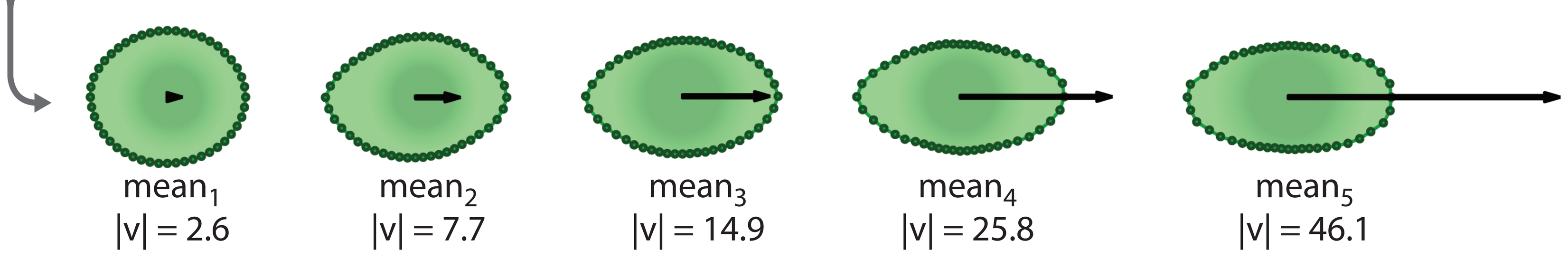
$$p(\mathbf{Z}_t \mid \mathbf{X}_t) = \prod_j p(Z_t^j|\mathbf{X}_t) = \prod_j p(Z_t^j \text{ is false alarm}) + \sum_i p(Z_t^j \mid X_t^i)\, p(X_t^i \text{ created } Z_t^j) + \text{higher order...}$$

$$p(Z_t^j \mid X_t^i) = p(L_t^j, A_t^j \mid M_t^i, O_t^i, R_t^i) \propto p(L_t^j \mid \mathrm{pos}(M_t^i))\ p(A_t^j \mid O_t^i)\ p(A_t^j \mid v(M_t^i))$$

Models the probability of an observed appearance given a motion hypothesis.

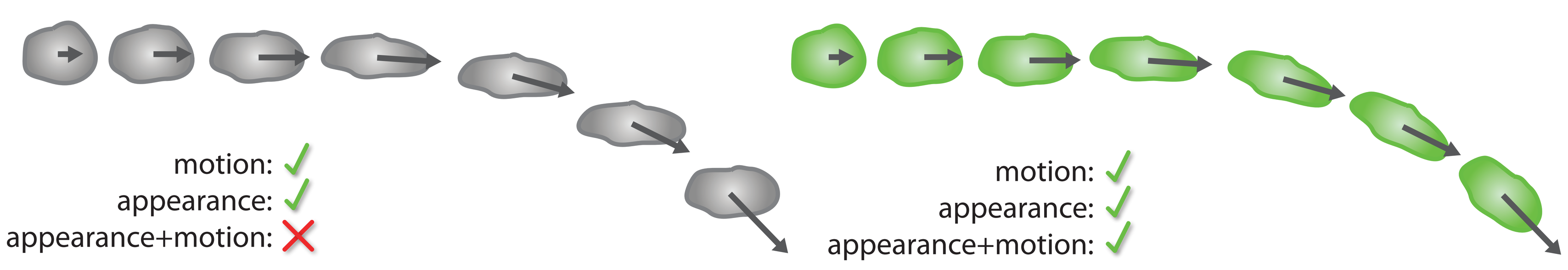
$p(A_t|v(M_t))$ is modeled as a 5-component GMM learned from nucleus motion and appearance training data.




(prototype nuclei constructed from centers of GMM components)

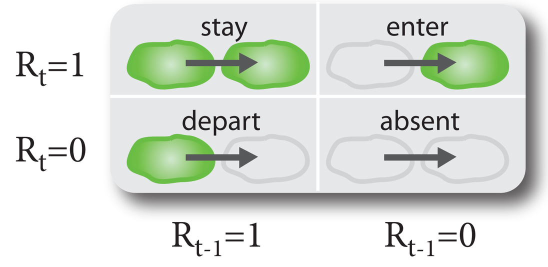
We search for good target paths by evaluating proposed paths according to the motion, appearance, and joint motion-appearance models.




For additional information & videos, visit http://cvlab.epfl.ch/~ksmith/

## The Second Constraint

**Constraint 2**: The entrance and departure of a target should occur near a boundary of the scene.




Presence is defined by $R_t$, $R_{t-1}$ for four cases.

The term $p(\mathbf{X}_t|\mathbf{X}_{t-1})$ is traditionally limited to a motion model. To enforce Constraint 2, $p(\mathbf{X}_t|\mathbf{X}_{t-1})$ also depends on the presence of targets at *t* and *t-1*.

$$p(\mathbf{X}_t|\mathbf{X}_{t-1}) = \prod_i p(X_t^i \mid X_{t-1}^i) = p(M_t^i, O_t^i, R_t^i \mid M_{t-1}^i, O_{t-1}^i, R_{t-1}^i)$$

$$= \underbrace{p(M_t^i \mid M_{t-1}^i, R_t^i, R_{t-1}^i)}_{\text{motion}}\ \underbrace{p(O_t^i \mid O_{t-1}^i, R_t^i, R_{t-1}^i)}_{\text{appearance}}\ \underbrace{p(R_t^i \mid M_{t-1}^i, R_{t-1}^i)}_{\text{presence}}$$

Motion:

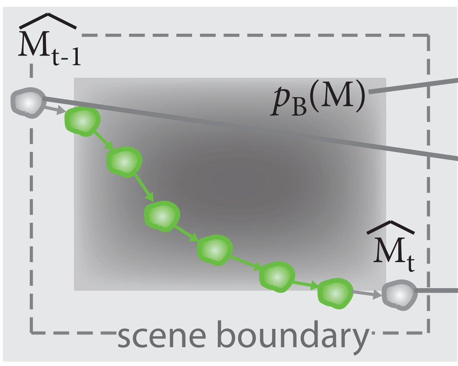
| | $R_{t-1}=1$ | $R_{t-1}=0$ |
|---|---|---|
| $R_t=1$ | $\mathcal{N}(M_t; \widehat{M}_t, \Sigma_M)$ | $p_B(\widehat{M}_{t-1})$ |
| $R_t=0$ | $\lambda_M$ | $\lambda_M$ |

Appearance:

| | $R_{t-1}=1$ | $R_{t-1}=0$ |
|---|---|---|
| $R_t=1$ | $\mathcal{N}(O_t; \widehat{O}_t, \Sigma_O)$ | $\mathcal{N}(O_t; \bar{O}, \Sigma_{\bar{O}})$ |
| $R_t=0$ | $\lambda_O$ | $\lambda_O$ |

Presence:

| | $R_{t-1}=1$ | $R_{t-1}=0$ |
|---|---|---|
| $R_t=1$ | $1\text{-}p_B(\widehat{M}_t)$ | $p_e$ |
| $R_t=0$ | $p_B(\widehat{M}_t)$ | $1\text{-}p_e$ |




$p_B(M)$ is a piecewise distribution modeling the probability of a target path entering or departing the scene.

The likelihood of the target path entering here is evaluated in the motion model for $R_{t-1}=0$, $R_t=1$ using back-prediction $\widehat{M}_{t-1}$, $p_B(\widehat{M}_{t-1})$.

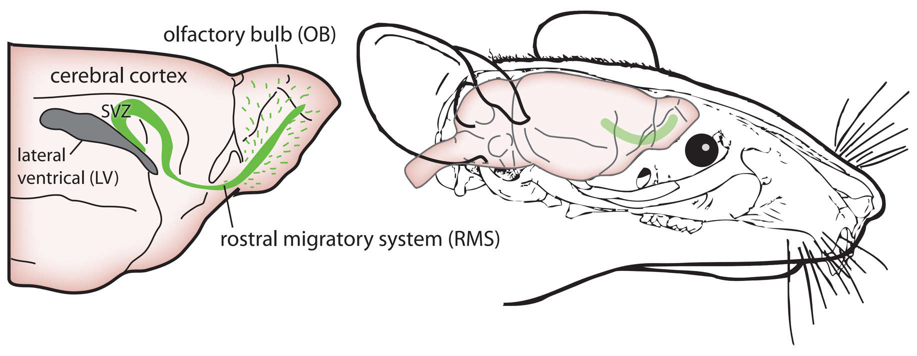
The likelihood of the target path exiting here is evaluated in the presence model for $R_{t-1}=1$, $R_t=0$ using the prediction $\widehat{M}_t$, $p_B(\widehat{M}_t)$.

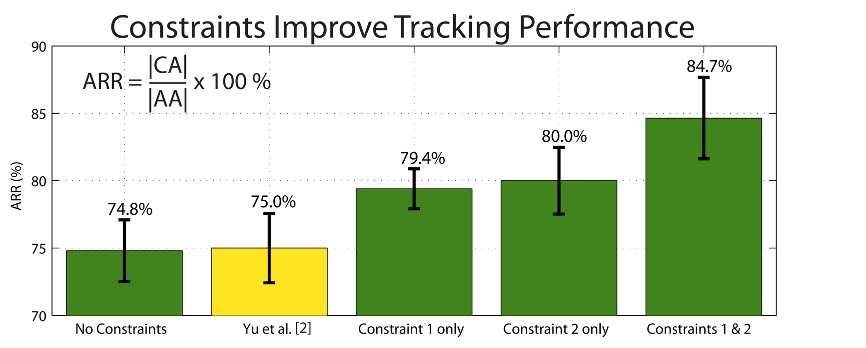
## Inference

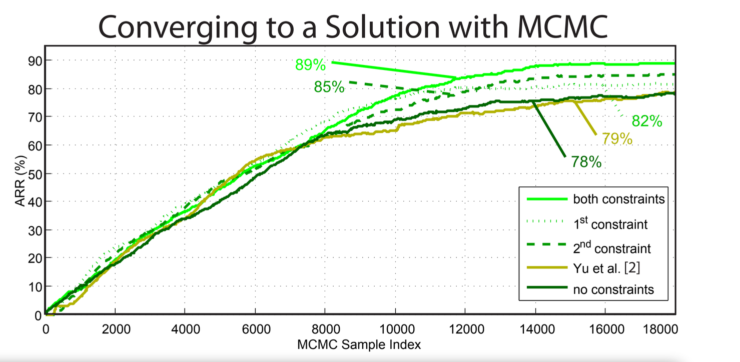
We use MCMC to efficiently estimate the MAP of (1). We initialize a Markov chain to an empty state and generate new samples by proposing changes to the previous state via a randomly selected MCMC move: *birth, death, associate, dissociate, merge, split*, or *swap* [1, 2]. The proposed state is added to the chain according to an acceptance probability, otherwise the previous state is added. After generating *N* samples, the MAP solution is given by the state with this highest posterior, $\widehat{\mathcal{X}}_T$.

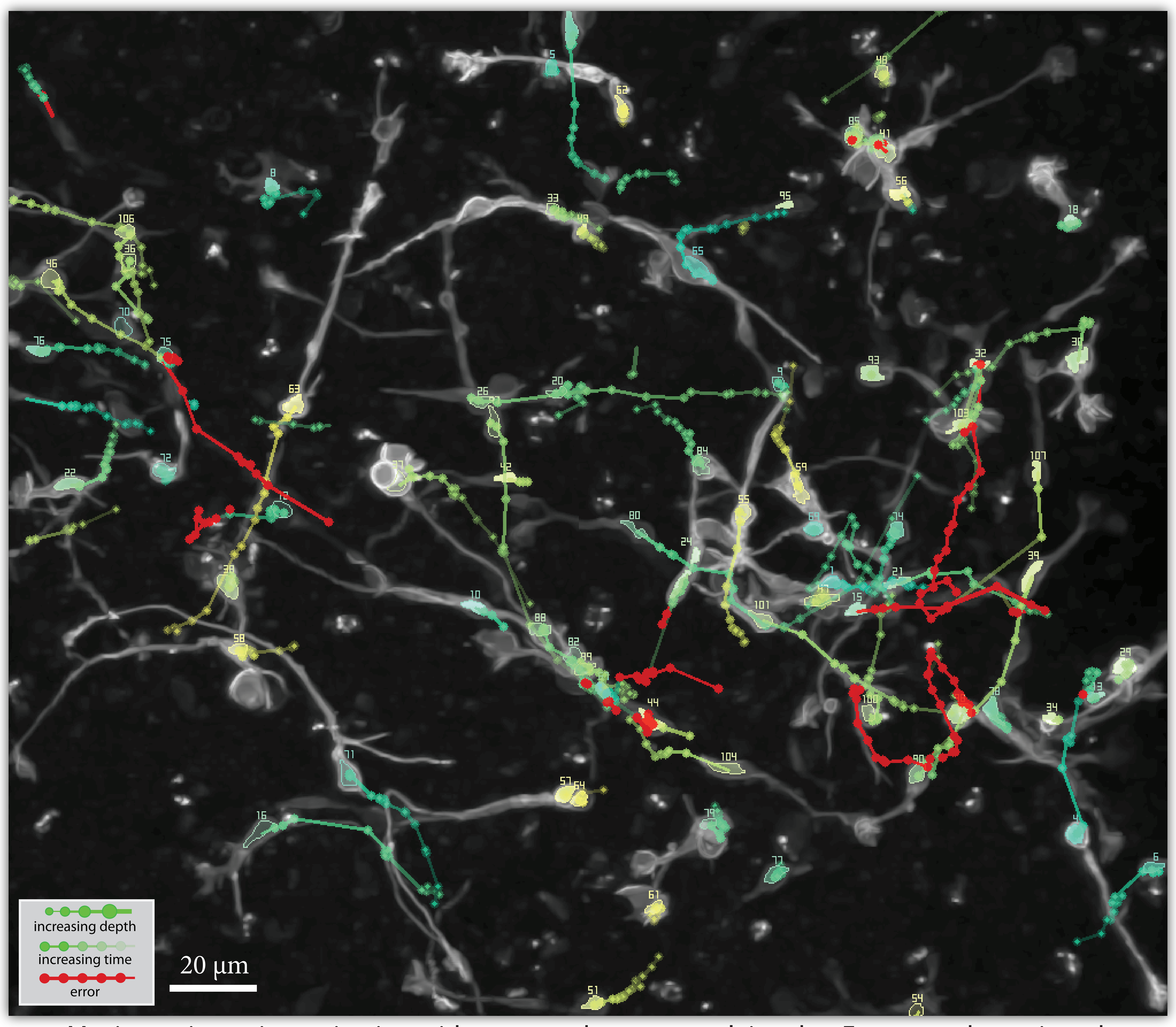
## Results for Neuron Videomicroscopy

We collaborate with neuroscientists studying neuroplasticity. A lentivirus is injected into the SVZ causing newly born neurons to express GFP. We image a 270 x 270 x 62 μm OB tissue sample with a 2-photon microscope.




Maximum intensity projection with recovered neuron nuclei paths. Errors are shown in red.


[1] S. Oh, S. Russell, and S. Sastry. Markov Chain Monte Carlo Data Association for General Multiple Target Tracking Problems. In *IEEE Conf. on Decision and Control*, 2004.

[2] Q. Yu, G. Medioni, and I. Cohen. Multiple Target Tracking Using Spatio-Temporal Markov Chain Monte Carlo Data Association. In *IEEE Conf. on Computer Vision and Pattern Recognition*, 2007.