Quantitative Analysis for Whole Brain Images of Caenorhabditis Elegans: a Machine Learning approach

Wu Stephen データ同化研究開発センター 特任助教

Yoshida, R. [1,2], Tokunaga, T.[2,3], Hirose, O.[1,2,4], Toyoshima, Y.[2,5], Teramoto, T.[2,6], Ishihara, T.[2,6], Iino, Y.[2,5]

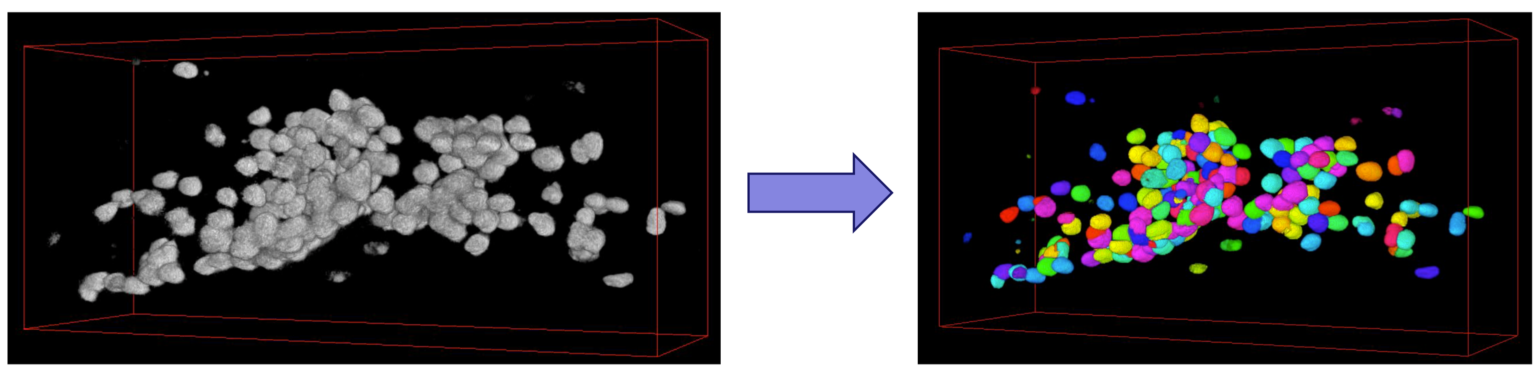
[1]The Institute of Statistical Mathematics, [2]JST-CREST, [3]Kyusyu Tech. Univ., [4]Kanazawa Univ., [5]Univ .of Tokyo, [6]Kyusyu Univ.

Abstract

A five-year-long cross-disciplinary project on studying whole brain neuronal activity of Caenorhabditis elegans (C. elegans) has taken part since 2013. The objective is to understand the underlying mechanisms of the complicated neural dynamics through recent advancements of 4D calcium imaging techniques. We tackle this great challenge with state-of-the-art machine learning techniques, which facilitates improved throughput of image processing. The machine learning pipeline begins with the detection and segmentation of imaged cells, followed by tracking of the crowded objects that exhibit great mobility in a time-lapse image sequence. For a given tracked cell, fluorescence intensities of the segmented voxels define temporal dynamics of its neural activities. Then, such information of all the brain cells is collectively analyzed to produce a network-based visualization, improving our understanding of the information processing mechanism of the neural system. This poster demonstrates the outline of the complete workflow.

(1) Tracking & (2) Segmentation

Gaussian Mixture Model + 3D Clump Splitting



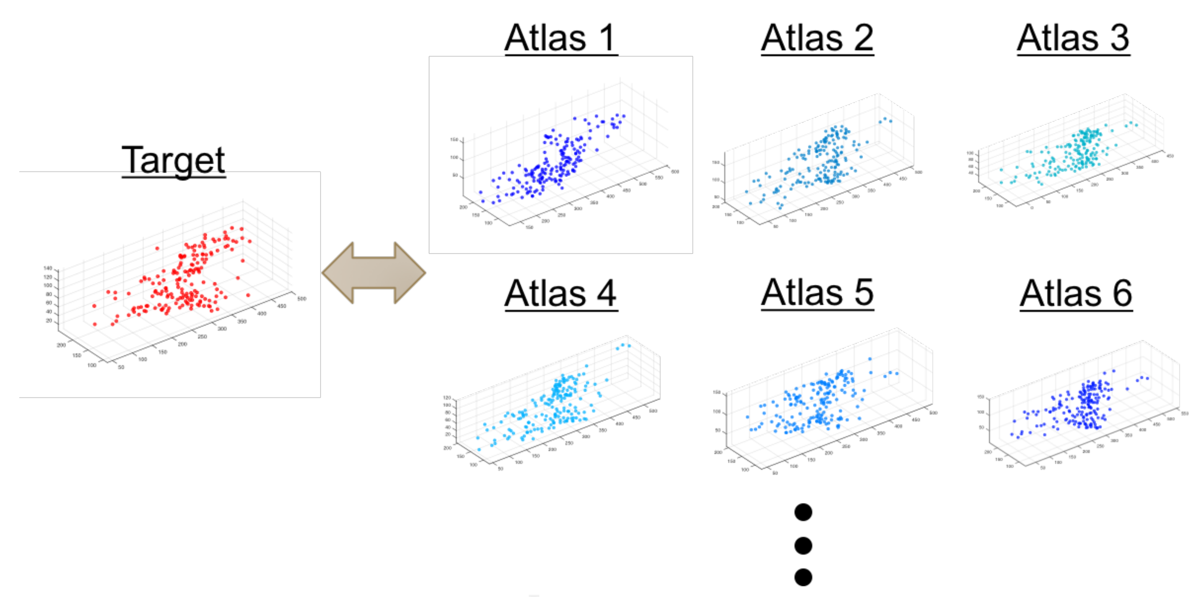
After standard de-noising and watershed segmentation steps, a special procedure is designed to separate overlapped neurons in the crowded area. As a result, a highly reliable automatic segmentation scheme is achieved that can identify some neurons missed by human expert segmentation.

Ref: Toyoshima et al. (2016), *PLOS Comp. Biol.* [in print]

(3) Annotation

Atlas-based Weighted Majority Voting

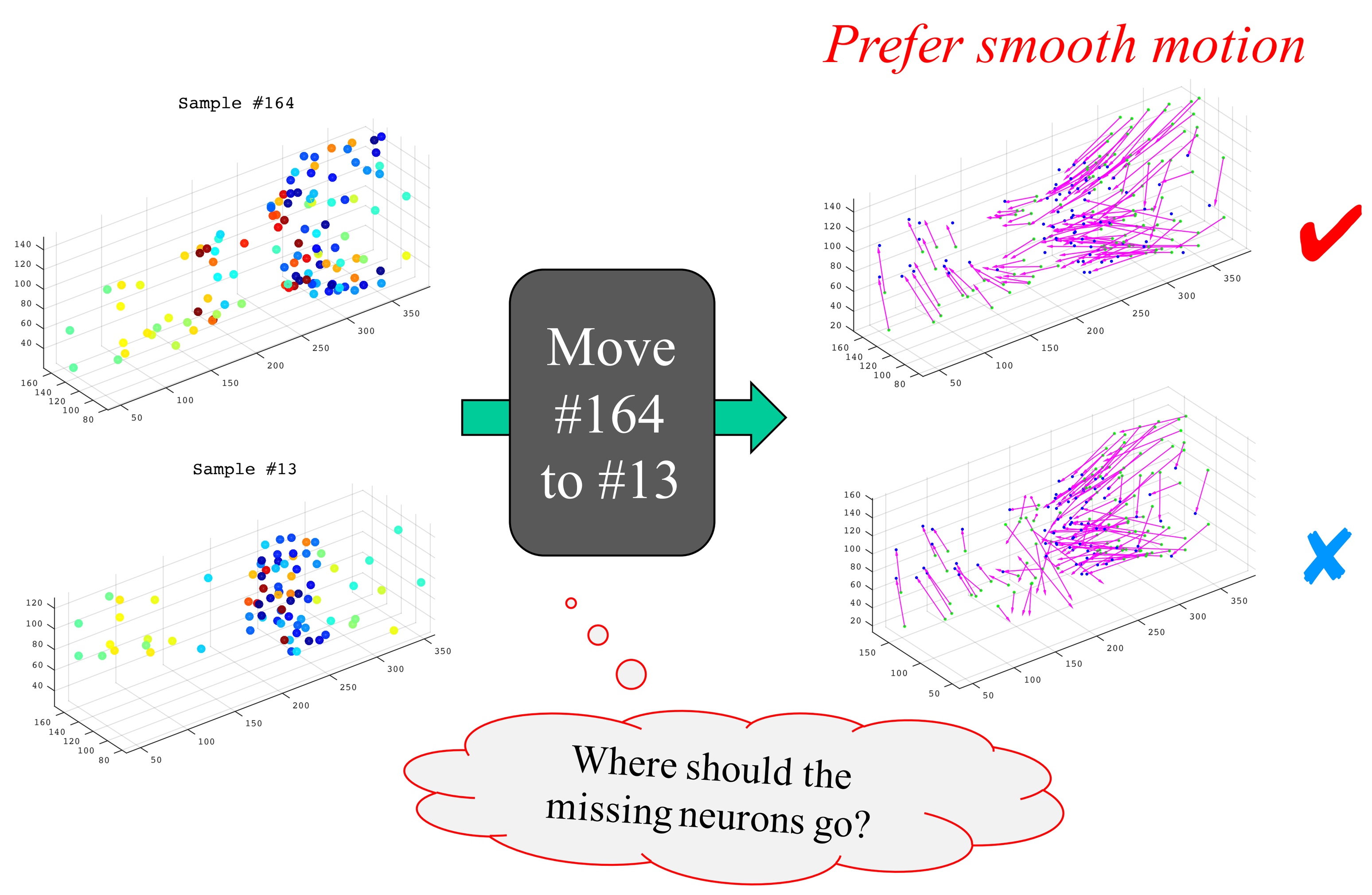
Match neurons from target to atlas → Weight the matching results based on confidence of accuracy → Combine results and find the best match



Target	Votes	
	AFDL	AINR
Neu. 1	10	124
Neu. 2	0	5
Neu. 3	43	28

(3a) Atlas Generation

Registration with Motion Coherent Theory



(3b) Bipartite Matching

1. Coherent Point Drift Registration

Move point set by minimizing energy: Ref: Myronenko & Song (2010)

$$E(\tilde{v}) = -\sum_{n=1}^{N} \log \sum_{m=1}^{M} e^{-\frac{1}{2}\left\| \frac{x_n - y_m}{\sigma} \right\|^2} + \frac{\lambda}{2} \int_{\mathbb{R}^d} \frac{|\tilde{v}(s)|^2}{\tilde{G}(s)} \, \mathrm{d}s$$

Gaussian Mixture Model → Robust registration

Suppress high frequency velocity field → Smoothness

2. Integer Linear Programming

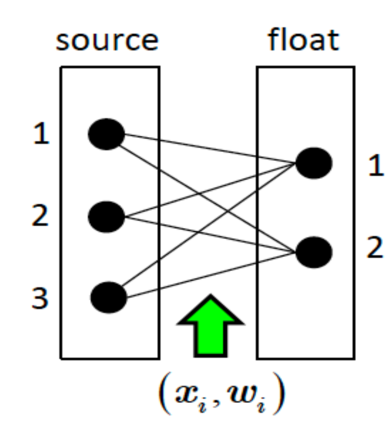
$\min_{x} \; w^{\mathrm{T}} x$

subject to

$x \in \{0,1\}^n$ Assignment variables (binaries)

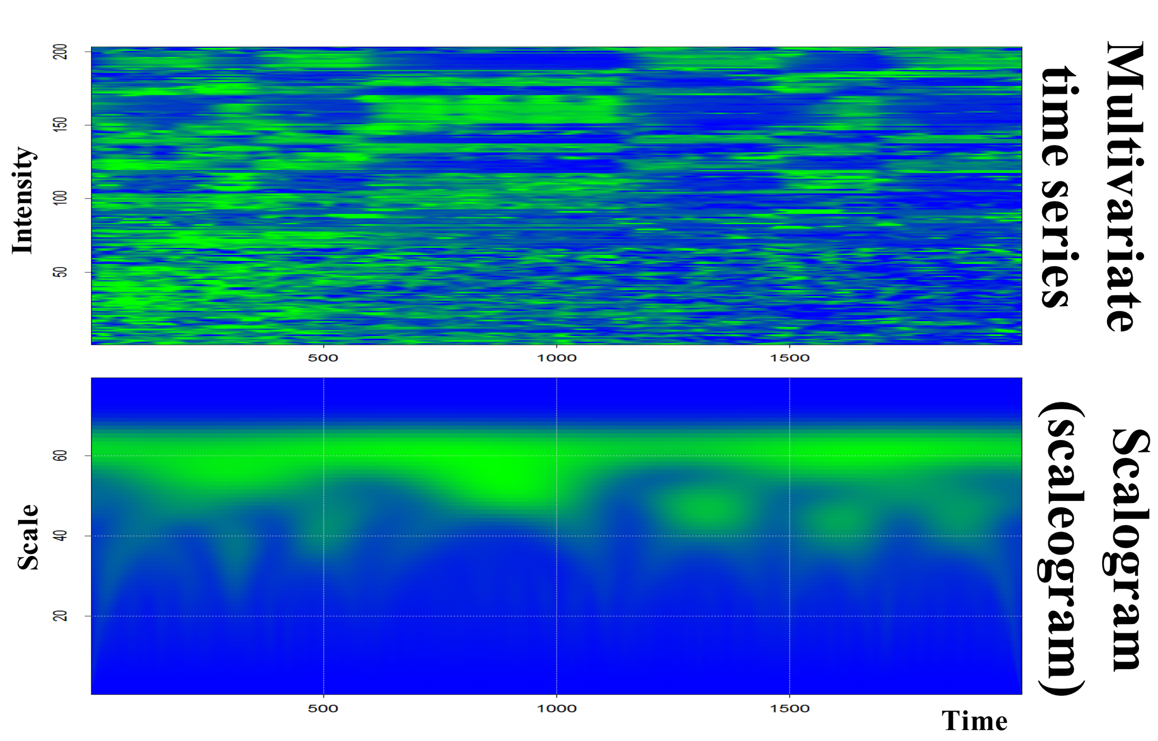
$Cx \le 1$ one-to-one matching

$1^{\mathrm{T}} x = \theta_{\max}$ Total # of assignments

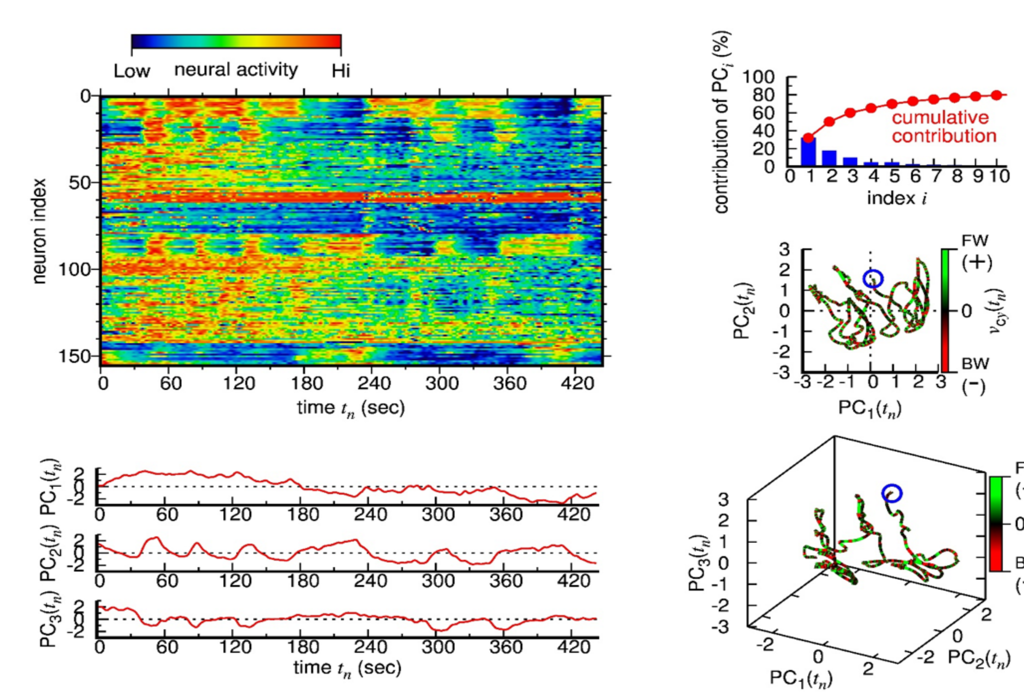


(4) Collective Data Analysis and Visualization

Annotation-free Analysis



Phase Clustering for PCA



Source: Y. Iwasaki, Ibaraki U. (College of Eng.)

Hive Plot Visualization