High-throughput Investigation of Drosophila Brains via Structure-Based Similarity

Florian Ganglberger, Laszlo Tirian, Florian Schulze, Andrew Straw and Katja Bühler

Identification of suitable transgenic driver lines for small subsets of Drosophila neurons is critical for understanding the role of single neurons and circuits of interconnected neurons in behavior. Comparing neuronal types in large datasets is a major challenge that is done typically by browsing through Z-projections with subsequent loading of the original confocal image. Besides its time-consuming nature, it prefers neurons that are not hidden behind other neurons, neurons with bright signal and neurons on images with only a few labeled neurons. Therefore, a high- throughput search in large databases requires automated tools for image retrieval.

We use a highly optimized structure based neuron retrieval method. Through sparsification, file compression, downscaling, multithreading and code optimizations we made fast and user-friendly investigation of big datasets possible. We tested the method on the Vienna Tile (VT) collection, a large publicly available registered 3D neuronal database (~11.000 images). We retrieved a large number of images representing several driver lines for all neuronal types we tested. Moreover, the method showed insensitivity to signal to noise ratio and expression broadness of the images. It perform robustly to biological variability and registration errors, since queries retrieved images with as little as 5% overlap of the queried voxels. After a preprocessing step, queries can be executed in ~20 sec per 1000 images.

In conclusion, this tool enables fast and comprehensive identification of suitable driver lines in large registered 3D image databases to support anatomical and/or behaviour studies or system biology approaches.