Who is talking to whom: synaptic partner detection in anisotropic volumes of insect brain

Anna Kreshuk[1,3], Jan Funke[2], Albert Cardona[3], and Fred A. Hamprecht[1]

HCI/IWR, University of Heidelberg
Institute of Neuroinformatics, UZH/ETH Zurich
HHMI Janelia Research Campus

Abstract. Automated reconstruction of neural connectivity graphs from electron microscopy image stacks is an essential step towards large-scale neural circuit mapping. While significant progress has recently been made in automated segmentation of neurons and detection of synapses, the problem of synaptic partner assignment for polyadic (one-to-many) synapses, prevalent in the *Drosophila* brain, remains unsolved. In this contribution, we propose a method which automatically assigns pre- and postsynaptic roles to neurites adjacent to a synaptic site. The method constructs a probabilistic graphical model over potential synaptic partner pairs which includes factors to account for a high rate of one-to-many connections, as well as the possibility of the same neuron to be pre-synaptic in one synapse and post-synaptic in another. The algorithm has been validated on a publicly available stack of ssTEM images of *Drosophila* neural tissue and has been shown to reconstruct most of the synaptic relations correctly.

Keywords: Circuit reconstruction, graphical model, electron microscopy

1 Introduction

Recent advances in electron microscopy instrumentation and sample preparation allow neuroscientists to acquire unprecedented volumes of data. Automating the analysis of the acquired images, however, still poses unsolved challenges to the computer vision community [7,14]. In order to reconstruct the connectivity in a block of neural tissue, neuroscientists have to trace all the neurons in the image stack and establish which of them are connected by synapses and in which direction. In the mammalian cortex, most synapses connect a single presynaptic neuron (sender) to a single postsynaptic neuron (receiver). Moreover, within the relatively small fields of view accessible to electron microscopy, mammalian cortex neurons mostly play an unambiguous sending or receiving role. Matters get more complicated in fruitfly neural tissue, where most synapses are polyadic (one-to-many, Fig. 1, Fig. 3), and, due to higher overall density of connections, more neurites are populated by both pre- and postsynaptic sites [3]. For such data, it becomes necessary to both find all the partners in a synaptic connection and to establish the direction of signal flow.

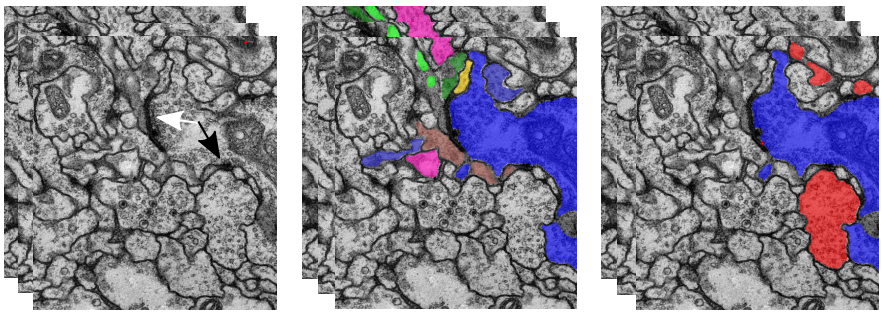


Fig. 1. A small substack of the data, showing the synaptic partners involved in two synapses. Left: raw data; black arrow points to a one-to-one synapse, white arrow to a one-to-many synapse from inside the presynaptic neuron. Center: pre- and postsynaptic neurons, participating in the larger one-to-many synapse (not all connections happen in the top section). Right: pre- and postsynaptic partners of the one-to-one synapse.

Until recently, all image analysis for neural circuit mapping has been performed manually. While manual processing continues to play a major role, latest advances in neuron segmentation and synapse detection algorithms allow neuroscientists to switch to a semi-automated mode with automated processing followed by targeted proofreading [10,17]. For FIB/SEM data with isotropic resolution [11], performance of automated synapse detection methods [2,13,18] is already comparable to that of human annotators. While anisotropic ssTEM data is more difficult, several methods have very recently been suggested to tackle this problem ([9,12,19]). The significantly more challenging problem of automated neuron segmentation has been the subject of even more active research. While human performance is not yet reached, substantial progress has been made in recent years [1,5,8,15,16,21]. Overshadowed by the challenges of neuron segmentation and synapse detection, the problem of synaptic partner assignment has so far been left to fully manual annotation. Automating this painstakingly laborious and error prone analysis step is the target of our contribution.

We propose to build a probabilistic graphical model, where synaptic roles of pairs of spatially adjacent neurites are modeled by random variables. Morphological properties of the neurons are incorporated into unary factors; general synapse properties, such as the prevalence of one-to-many connections, are modeled by pairwise factors. Our motivation for choosing this method is the simplicity of modeling rules of different strengths, from slight preference of certain neuron/synapse configurations to rules with almost no exceptions, such as neurons not synapsing on themselves. Pairwise factors allow us to model much more prior biological knowledge and substantially improve the results.

A detailed description of the model can be found in the next section. In section 3 we apply the method to a stack of *Drosophila* larva neural images with anisotropic resolution and demonstrate that this model is superior to neuron pair

classification. Finally, in section 4 we discuss how this model could be extended to isotropic data or used jointly with a neural segmentation procedure.

2 Methods

Our method starts from a segmentation of the stack into neurons and detection of its synapses. In the following we refer to individual images of the stack as sections and to cross sections of neurons in the image as neuron slices. Similar to neuron segmentation methods for anisotropic data, we consider the stack section by section and then introduce inter-section links.

In each section of the stack, we group adjacent neuron slices into pairs of neighbors. For a pair of neighbor neuron slices i and j we then introduce a random variable P_{ij} with three possible states: 1) no synaptic connection, 2) the first neuron of the pair is presynaptic, the second is postsynaptic, 3) the first neuron is postsynaptic, the second is presynaptic. Note that pre- and postsynaptic roles of the neuron slices are defined individually for each potential synaptic connection rather than globally. Thus, a neuron slice can be presynaptic to one and postsynaptic to another one of its neighbors.

2.1 Unary factors

The unary factors for each pair represent the prediction we can make on the state of the pair, when considering each pair individually. These are computed as predictions of a Random Forest classifier. The features for prediction are inspired by rules typically employed by human annotators, namely: 1) neurotransmitter vesicles are found on the presynaptic side of the connection; 2) presynaptic neuron slices are usually larger than postsynaptic ones; 3) there is a synaptic density (electron dense region, showing as a darker spot in the images) very near to the boundary between the slices. These cues are shown in Fig. 3(left). To compute the first feature, we train another Random Forest classifier to detect vesicles on the pixel level. The ilastik toolkit [20] has been used to train the classifier interactively on very sparse user labels. A similar classifier was used by [9] to detect vesicles in the retina, while [2] employ similar features to implicitly encode the presence of vesicles on one and absence on the other side of the synapse. The number of segmented vesicles is used as a feature for the unary Random Forest. To obtain an estimate of the synapse presence between the neurons of the pair, we either use smoothed precomputed synapse detections or, if these are not available, train another pixelwise classifier on synaptic densities. The summed predictions of this classifier in the vicinity of neuron-neuron boundary are also taken as a feature for the unary Random Forest. The remaining features are areas of the two neuron slices. The prediction of the Random Forest for the neuron slices i and j is then taken as the unary factor ψ_{ij} associated with the random variable P_{ij}.

2.2 Pairwise factors

The unary factors predict the synaptic relation for each pair of neuron slices individually. To reflect dependencies between pairs of neuron pairs, we augment our model by pairwise factors. In particular, we strive to encode the following biological knowledge: i) a single synapse should have only one presynaptic neuron (there are no convergent synapses in our test data); ii) one-to-many connections are more likely than one-to-one; iii) in cases when a neuron slice is both pre- and postsynaptic through different synapses, these synapses are usually not located immediately next to each other.

A pairwise factor is introduced for two random variables whenever their constituent pairs share one of the slices. For example, if P_{ij} represents the pair of slices i and j, and P_{ik} the pair of slices i and k, then we introduce a pairwise factor $\psi_{ij,ik}$ between them. The values of this factor depend not only on the states of P_{ij} and P_{ik}, but also on the conditions of them belonging to the same synapse and j and k being neighbors. Consequently, we introduce four variants of the $\psi_{ij,ik}$ factor with different value tables (see also Fig. 2):

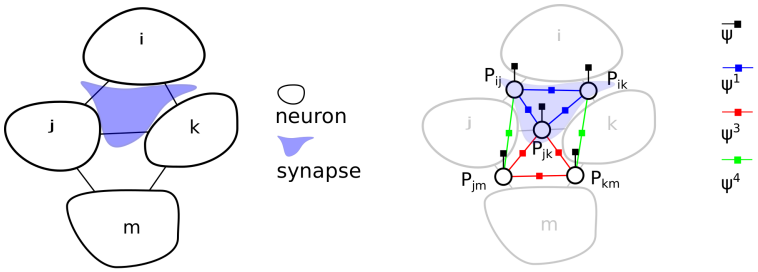
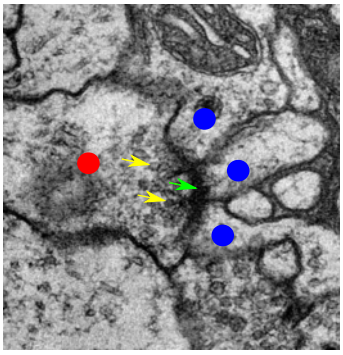


Fig. 2. An example of a region adjacency graph of neurons and synapses and the corresponding factor graph model. Left: a region adjacency graph for four neuron slices i, j, k, m and one synapse. Note that i and m are not neighbors, and the j to m and k to m boundaries are not covered by the synapse. Right: the corresponding factor graph for variables P_{ij}, P_{ik}, P_{jk}, P_{jm}, P_{km}. Connections across different sections of the stack are omitted. See main text in Section 2.2 for the factor legend.

1. $\psi^1_{ij,ik}$: j and k are neighbors, i, j and k touch the same synapse. In this case, we penalize the variable states which give rise to the configurations with two presynaptic neurons in the same synapse (cost c_0), neuron being pre- and postsynaptic through the same synapse (cost c_1) and one-to-one rather than one-to-many connection (cost c_2). This potential, along with the corresponding variable states, is shown in Fig. 3(right).
2. $\psi^2_{ij,ik}$: j and k are *not* neighbors, i, j and k touch the same synapse. In this case we again penalize variable states which lead to two presynaptic neurons in the same synapse and a neuron being both pre- and postsynaptic through the same synapse.

3. $\psi^3_{ij,ik}$: j and k are neighbors, i, j and k do *not* touch the same synapse. Here we penalize the state with a many-to-one connection, as well as the state with one neuron being both pre- and post-synaptic, but with a lower cost than for neurons touching the same synapse.
4. $\psi^4_{ij,ik}$ (j and k are *not* neighbors, i, j and k do *not* touch the same synapse. In this case we slightly penalize the state, where one neuron is both pre- and postsynaptic.

Fig. 3 shows the values of $\psi^1_{ij,ik}$ as a function of the state of P_{ij} and P_{ik}.



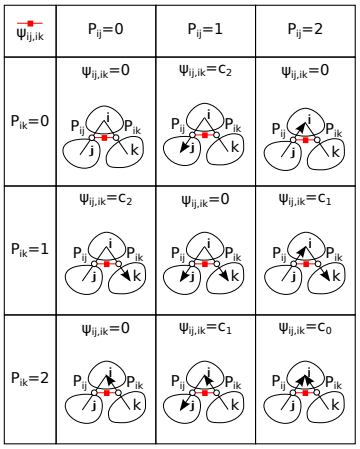


Fig. 3. Left: a typical one-to-many synapse. Red dot marks the pre-synaptic neuron, blue dots its post-synaptic partners. Yellow arrows point to vesicles and the green arrow to the electron-dense area of the membrane. Right: the value table for the $\psi^1_{ij,ik}$ factor for different values of P_{ij} and P_{ik}. Random variable states are denoted by arrows showing synapse direction: $P_{ij} = 0$ – no synaptic connection, $P_{ij} = 1$ – i is presynaptic, j is postsynaptic, $P_{ij} = 2$ – vice versa.

To link the variables between consecutive sections, we additionally introduce pairwise factors between pairs of slices of the same neurons which encourage consistency of the synaptic role assignment between sections. For illustration, let neuron n_i be represented by slice i_s in section s and by slice i_{s+1} in section $s+1$, while neuron n_j is represented by j_s and j_{s+1} respectively. If both i_s, j_s and i_{s+1}, j_{s+1} form neighor pairs, touching the same synapse in both sections, we add a pairwise factor between variables $P_{i_s,j_s}, P_{i_{s+1},j_{s+1}}$.

The MAP estimation problem for this factor graph can be reformulated as an energy minimization problem, with the usual transition from potentials to energies ($\mathcal{N}$ stands for the set of all pairs of neighbors):

$$\psi_{ij}(P_{ij}) = \exp(-E_{ij}(P_{ij})), \qquad \psi_{ij,ik}(P_{ij}, P_{ik}) = \exp(-E_{ij,ik}(P_{ij}, P_{ik}))$$

$$\underset{P_{ij},(i,j)\in\mathcal{N}}{\operatorname{argmin}} E(P) = \underset{P_{ij},(i,j)\in\mathcal{N}}{\operatorname{argmin}} \sum_{(i,j)\in\mathcal{N}} E_{ij} + \sum_{(i,j),(i,k)\in\mathcal{N}} E_{ij,ik}$$

We solve this problem to global optimality by an ILP solver.

3 Results

The proposed algorithm has been validated on a publicly available dataset from *Drosophila* larval neuropil volume [6]. Two 1024×1024×20 image stacks are provided along with manual neuron segmentation and synapse detections for the first stack. We annotated the synapses in the second stack and the pre- and post-synaptic neuron slices in both stacks. The first stack has been used for training the unary Random Forest and the first 9 sections of the second stack have been used for the grid search of the best values for the costs in the pairwise potentials.

Table 1. Algorithm results. First and second column show recall and precision of the synaptic partner detection. The third column shows the total number of synaptic partner pairs in the groundtruth annotation. Fourth column shows the number of synapses, for which all partners have been detected automatically (with possible false positive pairs besides true detections). Fifth and sixth columns show the number of synapses, where not all or no partners at all have been found. See main text for evaluation details.

Method	Recall	Precision	Total pairs	Fully recovered synapse groups	Partially recovered synapse groups	Not recovered synapse groups
Unary only	90%	64%	163	36	6	5
Full model	90%	78%	163	38	2	7

The evaluation procedure ran as follows: groundtruth neuron segmentation and synapse detection were provided as input to the algorithm. Each synaptic partner pair, detected by the algorithm, was assigned to its closest synapse, if it was within a 10 pixel distance from it on any section. Pairs with no synapse at such distance were discarded. Recall was computed as the number of groundtruth pairs found by the algorithm, divided by the total number of groundtruth pairs; precision as the number of algorithm pairs found in the groundtruth, divided by the total number of algorithm pairs. A synapse group is "fully recovered"/"partially recovered", if all/at least one of its pairs are found. The groundtruth contained 47 synapses, connecting 163 neuron pairs. For 7 such synapses, the algorithm did not find any of the pairs. We analyzed these synapses in more detail and found that 4 out of 7 were small synapses, only visible in the first or last slice. Another missed synapse was located very close to the image border. A difficult missed connection of a partially recovered synapse group is shown in Fig. 4(A, B, C, D). Other connections of this synapse, found in consecutive sections, were reconstructed correctly. A typical false positive, caused by the preference for the same label for neighboring pairs, is shown in Fig. 4(E,F). The same preference explains the reduced number of partially recovered synapse groups for the full model in Table 1.

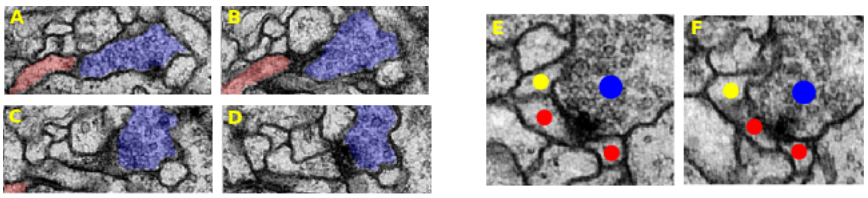


Fig. 4. A, B, C, D: serial sections of a missed connection. Presynaptic slice is colored blue, postsynaptic red. E, F: a typical false positive error over two sections. Blue dot: presynaptic neuron, red dots: correct postsynaptic neurons, yellow dot: false positive.

4 Discussion

We have introduced an algorithm for synaptic partner detection, which, by constructing a probabilistic graphical model to couple synaptic partner pairs, can benefit from more biological prior knowledge than an estimator considering each pair individually. On a test dataset, our evaluation shows a significant improvement of precision without a loss in recall.

In its current form, our algorithm explicitly uses 2D neuron slices and links them together across the lower resolution z-axis. It could, however, also be extended to isotropic 3D data, such as data produced by FIB/SEM microscopes. For a fully 3D approach, we would consider neuron pairs within a spherical neighborhood of each synapse detection and replace the features by their 3D counterparts.

Assignment of synaptic partners could also be incorporated directly into the neuron segmentation procedure and perhaps improve both the segmentation and the partner assignment steps. For methods operating on neuron hypotheses [4,15,21], this could be achieved by explicitly introducing synaptic relations between neuron slice candidates. Segmentation could then make use not only of the general prior knowledge of neuron appearance, but also more specific local pre- and postsynaptic features.

We believe that the proposed algorithm can substantially reduce human annotation effort for insect brain circuit reconstruction by shifting the synaptic partner detection task from fully manual to at least semi-automated domain, as well as pointing out the human attention errors.

References

1. Andres, B., Kroeger, T., Briggman, K.L., Denk, W., Korogod, N., Knott, G., Koethe, U., Hamprecht, F.A.: Globally Optimal Closed-surface Segmentation for Connectomics. Proceedings of ECCV (2012)
2. Becker, C., Ali, K., Knott, G., Fua, P.: Learning Context Cues for Synapse Segmentation. Proceedings of MICCAI (2012)
3. Cardona, A., Saalfeld, S., Preibisch, S., Schmid, B., Cheng, A., Pulokas, J., Tomancak, P., Hartenstein, V.: An integrated micro- and macroarchitectural analysis of

the Drosophila brain by computer-assisted serial section electron microscopy. PLoS biology 8(10) (Jan 2010)
4. Funke, J., Andres, B., Hamprecht, F.A., Cardona, A., Cook, M.: Efficient automatic 3D-reconstruction of branching neurons from EM data. Proceedings of CVPR (2012)
5. Funke, J., Martel, J., Gerhard, S., Andres, B., Ciresan, D.C., Giusti, A., Gambardella, L.M., Schmidhuber, J., Pfister, H., Cardona, A., Cook, M.: Candidate Sampling for Neuron Reconstruction from Anisotropic Electron Microscopy Volumes. In: Proceedings of MICCAI (2014)
6. Gerhard, S., Funke, J., Julien, M., Cardona, A., Fetter, R.: Segmented anisotropic ssTEM dataset of neural tissue. `http://dx.doi.org/10.6084/m9.figshare.856713`
7. Helmstaedter, M., Mitra, P.P.: Computational methods and challenges for large-scale circuit mapping. Current opinion in neurobiology 22(1), 162–169 (2012)
8. Huang, G.B., Jain, V.: Deep and Wide Multiscale Recursive Networks for Robust Image Labeling. Proceedings of ICLR (2014)
9. Jagadeesh, V., Anderson, J., Jones, B., Marc, R., Fisher, S., Manjunath, B.S.: Synapse Classification and Localization in Electron Micrographs. Pattern Recognition Letters (2013)
10. Kaynig, V., Knowles-barley, S., Jones, T.R.: Large-Scale Automatic Reconstruction of Neuronal Processes from Electron Microscopy Images. arxiv:1303.7186v1 (2013)
11. Knott, G., Marchman, H., Wall, D., Lich, B.: Serial section scanning electron microscopy of adult brain tissue using focused ion beam milling. The Journal of neuroscience 28(12), 2959–2964 (2008)
12. Kreshuk, A., Köthe, U., Pax, E., Bock, D.D., Hamprecht, F.A.: Automated Detection of Synapses in Serial Section Transmission Electron Microscopy Image Stacks. PLoS ONE 9, 2 (2014)
13. Kreshuk, A., Straehle, C.N., Sommer, C., Koethe, U., Cantoni, M., Knott, G., Hamprecht, F.A.: Automated detection and segmentation of synaptic contacts in nearly isotropic serial electron microscopy images. PloS ONE 6(10), e24899 (2011)
14. Lichtman, J.W., Pfister, H., Shavit, N.: The big data challenges of connectomics. Nature Neuroscience 17(11), 1448–1454 (Oct 2014)
15. Liu, T., Jones, C., Seyedhosseini, M., Tasdizen, T.: A modular hierarchical approach to 3D electron microscopy image segmentation. Journal of neuroscience methods 226, 88–102 (Apr 2014)
16. Nunez-Iglesias, J., Kennedy, R., Parag, T., Shi, J., Chklovskii, D.B.: Machine Learning of Hierarchical Clustering to Segment 2D and 3D Images. PloS one 8(8), e71715 (Jan 2013)
17. Plaza, S.M.: Focused Proofreading : Efficiently Extracting Connectomes from Segmented EM Images. arXiv:1409.1199v1 (2014)
18. Plaza, S.M., Parag, T., Huang, G.B., Olbris, D.J., Saunders, M.A., Rivlin, P.K.: Annotating Synapses in Large EM Datasets. arXiv:1409.1801v2 (2014)
19. Roncal, W.G., Kaynig-Fittkau, V., Kasthuri, N., Berger, D., Vogelstein, J.T., Fernandez, L.R., Lichtman, J.W., Vogelstein, R.J., Pfister, H., Hager, G.D.: Volumetric Exploitation of Synaptic Information using Context Localization and Evaluation. arXiv:1403.3724 (2014)
20. Sommer, C., Straehle, C., Kothe, U., Hamprecht, F.A.: Ilastik: Interactive learning and segmentation toolkit. Proceedings of ISBI (2011)
21. Vazquez-Reina, A., Gelbart, M., Huang, D., Lichtman, J., Miller, E., Pfister, H.: Segmentation fusion for connectomics. Proceedings of ICCV (2011)