Construction of Anatomically Correct Models of Mouse Brain Networks

Bruce McCormick, Yoonsuck Choe, and Wonryull Koh
Department of Computer Science, Texas A&M University, College Station, TX 77843
{mccormick,wkoh,choe}@cs.tamu.edu

Abstract

The *Mouse Brain Web*, a federated database, provides for the construction of anatomically correct reconstructions of the entire mouse brain network. Each web page in this database will provide the position, orientation, morphology, and synapses for each biologically observed neuron. We sketch here the components for constructing these networks, specifically (1) the spatial distribution of neurons by type; (2) wiring the network – synaptic assembly; (3) projection of neuron morphology and synapses to geometric multi-compartmental models; (4) using customized web-crawlers on the *Mouse Brain Web* to search out motifs and canonical circuits in the networks; and (5) the mapping of anatomically correct networks to physiologically correct network simulations.

Keywords: *Mouse brain networks, web-based database, neuron morphologies, synaptic assembly, network motifs.*

1 The *Mouse Brain Web*

The *Mouse Brain Web* database consists of interconnected webpages where a webpage represents an observed neuron and the hyperlinks between pages represent synapses between neurons. Each neuron in the Mouse Brain Web includes both its morphological and geometric data and its connectivity data extracted from our microstructure database [13].

We have designed a brain microstructure database system [13] to support 3D reconstruction and modeling of neurons and their networks from the microstructure data produced by three-dimensional microscopy: from the Brain Tissue Scanner (BTS; [17, 18]), confocal microscopy (CFM), and/or multiphoton microscopy (MPM). The brain microstructure database system contains two types of data that can be used to build the Mouse Brain Web: individual neurons (segment/neuron data) and their connectivity (brain network data). The segment/neuron data describes the geometry and morphology of each neuron and its subsequent, inferred type, represented by a set of attributes [7, 13, 19] defined in the database system. In addition, we have designed XML DTDs (Document Type Definitions) to map the morphology and synaptic data of each neuron in our object-relational database into XML files. These XML files are then converted to HTML and organized in the Mouse Brain Web.

The data size required to describe the geometry and morphology (segment/neuron data) of each neuron in the Mouse Brain Web depends on its type (which determines the approximate number of processes emanating from the soma and whether the dendritic processes have spines) and the depth of its axonal/dendritic processes (which determine the number of segments in each process). In addition to its geometry and morphology, a neuron in the Mouse Brain Web needs to contain its connectivity data extracted from the observed synapses (brain network data) as hyperlinks to its pre- and post-synaptic neurons. Each neuron in mammalian cortex is pre-synaptic to 7,000–8,000 neurons and post-synaptic to 6,000–10,000 neurons, and multiple synapses between the same two neurons are rare [4]. Following the statistics reported in the literature [4, 22, 29], we estimate that the data storage required to record a neuron's geometry, morphology and connectivity data plus its markup overhead is 1.04 megabytes. The storage requirement for the Mouse Brain Web, if all of 17 million neurons in the mouse cortex are to be linked together, then is $\sim$18 terabytes. Extending the *Mouse Brain Web* from the cerebral cortex (16M neurons) to the full brain (75M neurons) entails approximately 5-fold increase in storage requirements, or 90TB.

Our initial data to assemble the Mouse Brain Web, however, will come from transgenic GAT1-GFP mice. Therefore, we expect to observe only 15% of total neurons present in the mouse cortex (i.e., all GABA-ergic interneurons). Consequently, each observed synapse may not have an observed pre- or post-synaptic neuron, reducing further the storage requirement for the GAT1-GFP mouse brains.

2 Spatial distribution of neurons by type

The counts, spatial distribution, and morphology of cell bodies will be derived from scanning Nissl- or ethidium bromide-stained tissue with a 10X objective. The 2.5mm field of view of this objective will be imaged at 2048 pixels/scan line, or 1.2μm sampling interval. This resolution allows us to reconstruct the individual cell bodies in 3D, and avoids many of the ambiguities of stereological measurements.

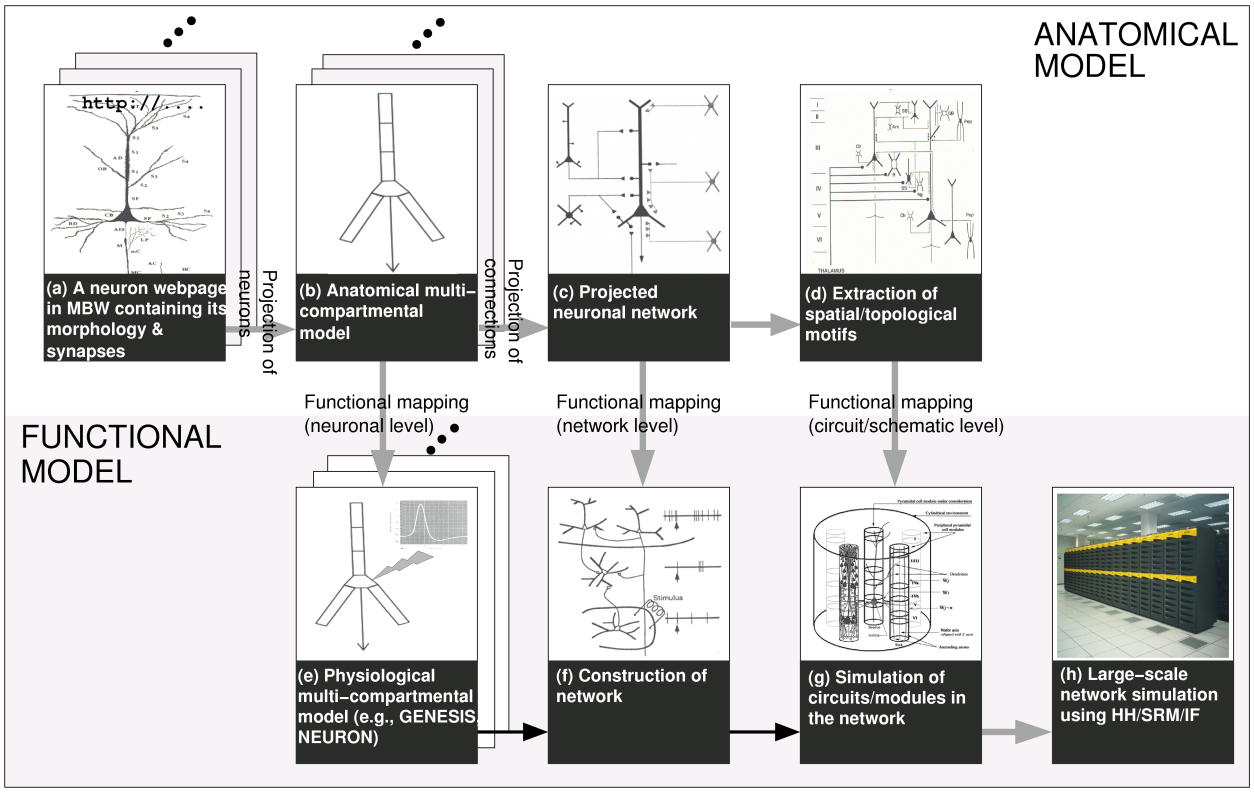

Figure 1: **Anatomically correct models of mouse brain networks and their functional simulation.** (a) and (d) are adapted from [22]; (b) and (e) are adapted from [3]; (c) and (f) are adapted from [29]; (h) See [26].

Cortical layer and fiber system identification. Generally the neuroanatomist determines, a priori, the boundaries of the six cortical layers, or lamina. Instead we will allow the mouse cortex data to "speak for themselves," through statistical classification (spatial partitioning) of the somata and fiber densities. Laminar boundaries are then mapped back to the 3D mouse brain atlas (Figure 2). As laminar boundaries are fuzzy, we will try the "method of mixtures" with fuzzy functions, to assign cells to each cortical layer. We encode the six layers in mouse cortex as I, II, III, IV, V, and VI, and the two fiber systems as S (Subcortical) and V (Vertical) (for positively stained fibers using osmium tetroxide, Brodal silver protrargol (unmyelinated) and Weigert methods).

Preliminary classification by cell type. With the criteria developed in step 1 above, we will sort the reconstructed neurons of each dendritic-stained tissue block by laminar layer. Then we classify individual neurons provisionally by cell type (e.g., pyramidal cells, spiny stellates, bitufted, chandelier, etc.), using the back-propagation algorithm and related supervised learning techniques of artificial neural networks.

Neuron subtype identification. At this point we may classify neurons by laminar level, soma morphology, and dendritic morphology.

Transgenic mice will reveal the cell type of fluorescing neurons directly: "In the hippocampus, virtually all GABAergic neurons, and only GABAergic neurons, show robust mGAT1-GFP fluorescence" (Chiu et al. [5]). This advantage greatly sharpens the classification of neurons.

Transgenic GAT1-GFP mice will also receive injections of retrograde and/or anterograde fluorescent tracers into major input-output structures of the cortex, such as thalamus, other cortices, striatum, and aminergic brainstem. This will allow the reconstruction of (1) locally connected GABA interneurons (GAT1-GFP labeled), (2) retrograde-labeled pyramidal neurons (one color retrograde tracer, e.g., Fluoro-Gold), and (3) anterograde-labeled inputs (with dextran/Pha-l). Thus, the local-circuit neurons (GABA-ergic) and the major output neurons (pyramidal) of the cortical module will be defined anatomically with other structures related functionally to the cortical module.

3 Wiring the network – Synaptic assembly

Neurons communicate at synapses. Recently the synaptic structure of GFP-filled neurons has been exhibited clearly by Chiu et al. [5] using confocal microscopy with a 100X objective (NA 1.4), and direct images of synapses like these

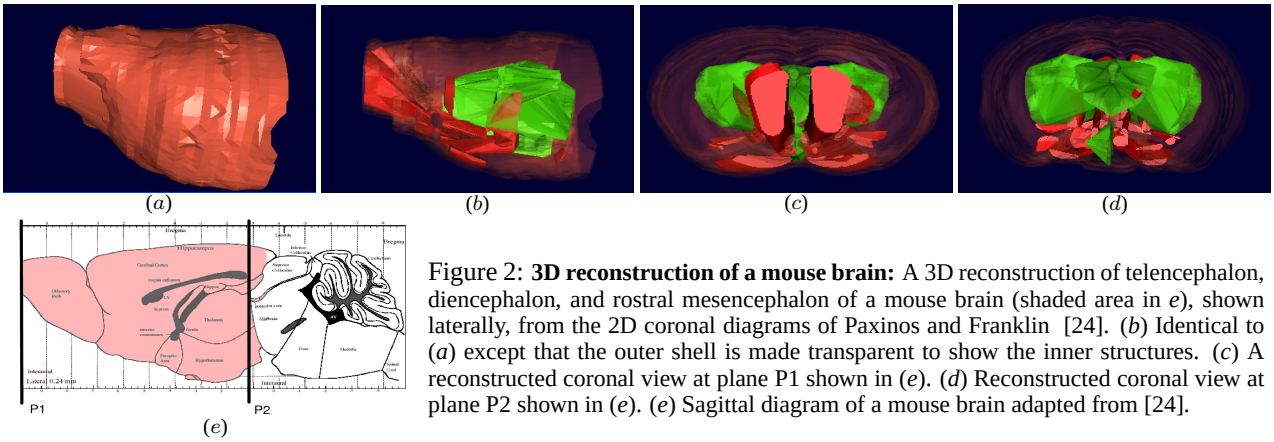

Figure 2: **3D reconstruction of a mouse brain:** A 3D reconstruction of telencephalon, diencephalon, and rostral mesencephalon of a mouse brain (shaded area in *e*), shown laterally, from the 2D coronal diagrams of Paxinos and Franklin [24]. (*b*) Identical to (*a*) except that the outer shell is made transparent to show the inner structures. (*c*) A reconstructed coronal view at plane P1 shown in (*e*). (*d*) Reconstructed coronal view at plane P2 shown in (*e*). (*e*) Sagittal diagram of a mouse brain adapted from [24].

will certainly improve under high-resolution MPM.

For each reconstructed neuron and for each of its identified postsynaptic terminals,we will identify, where possible, the presynaptic fibers, which are usually axons. The neuronal axons make a presynaptic contact with a postsynaptic terminal on the reconstructed neuron. In general, axons making presynaptic contacts extend from neuronal cell bodies located outside the volume under consideration. As synaptic assembly of neurons into a brain network is computationally very expensive, we will focus on selected regions of the mouse brain, including a significant portion of the cerebral cortex. We use two strategies to label putative synapses: geometric *proximity synapse labeling*, and *biochemical synapse labeling*.

Proximity synapse labeling. We will begin the assembly of the brain network by identifying potential synapses. We will first quickly identify pairs of "nearby" neurons that potentially form synapses, performing repeated proximity tests to identify all pairs of segments of the desired type within some threshold distance, typically 2–4μm [16]. We then employ a series of increasingly precise proximity tests, based on known or measured synaptic affinities [14], to refine the set of connectable pairs of neurons.

Biochemical synapse labeling follows complementary approaches: (1) Fluorescence-level measurement of GFP-expressed GABA transporter subtype 1 (GAT1) molecules for identifying inhibitory synapses. To measure GAT1 density precisely, we calibrate against known GFP density in transparent beads [5]. (2) Immunocytochemical staining of blocks of mouse brain tissue will select antibodies directed against specific presynaptic markers, such as synaptophysin, synaptobrevin, or SNAP 25. For identifying the synaptic complex, PZD 93 and PSD 95 may be used separately or in combination. The latter recognizes the excitatory glutamatergic synapses. The combination of these stains and GFP creates ideal synapse markers for MPM and the BTS. To aid comparison with previous studies, methods for the determination of synaptic density described by Braitenberg and Schuz for EM [4], and Chiu et al. [5] for MPM will be followed closely.

Database of synaptic counts. Neuroanatomical studies will independently validate synaptic counts derived from the BTS analyses with those from the MPM level [5]. These synaptic measures will be compared for separate subcompartments of the neuron, including dendrites, spines, somata, axon hillocks, and axon branches and terminals. We will create a database of synaptic counts for each of the cutting, staining, and reconstruction techniques used at the MPM and BTS levels of resolution, and will compare these counts with estimates in the literature [5, 9]. These methods are complementary, and creating a universal database will assist researchers with comparisons using both similar and different stains and visualization methods.

4 Projection of neuron morphology and synapses to multi-compartmental models

Geometrically reconstructed neurons (Section 2) and their connectivity (Section 3) will provide a data set that can be easily mapped to standard multi-compartmental models [11, 27] (Figure 3). The *Mouse Brain Web* database (Section 1) makes the mapping process significantly easier: The XML representations of the *Mouse Brain Web* provide a machine parsable way of specifying cell morphology for the generation of multi-compartmental simulation scripts.

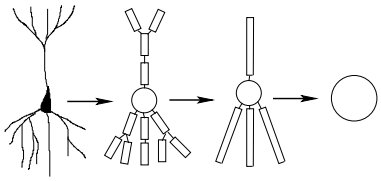

Figure 3: **Neuron data to multi-compartments.** Adapted from [6].

We will map each soma, axon, and dendritic segment in the geometric model into a single compartment in the multi-compartment model (one-to-one map-

ping). In cases where computational efficiency is critical, we will map an aggregate of several segments into an equivalent single compartment (many-to-one mapping). The resulting multi-compartmental model will be used in functional simulations (Figure 1).

5 Motif analysis

The hyperlink structure of HTML represents connectivity graphs in a natural way. To gather motifs and connection statistics, we will modify open-source web spiders (e.g., Grub [1]) to perform hyperlink analysis and site indexing. Our first spider implementation will be able to: (1) Index links at higher-order levels to deduce indirect connectivity and cycles; (2) Index motifs; (3) Generate repositories of connectivity statistics; and (4) Estimate velocity of neuronal spike conduction from axon morphology.

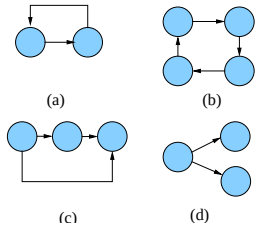

Figure 4: **Common motifs in complex networks.** (a) Autoregulation. (b) Multi-component loop. (c) Feedforward loop. (d) Single input motif. Adapted from [15].

For connectivity analysis and statistics calculation, we will adopt the graph analysis strategy developed by Sporns, et al. [30, 31][2] and incorporate the modified analysis engine into our crawler. The resulting spider and indexing engine, with the *Mouse Brain Web* database it accesses, will serve as a local and global resource for all research groups.

To analyze the *Mouse Brain Web* systematically, we rely on two aspects of the connectivity structure in the *Mouse Brain Web*: (1) a *neuronal connectivity graph*, and (2) *axonal morphology data* based on fiber length, diameter, and degree of myelination.

1. The *neuronal connectivity graph* will reveal motifs and modules (Figure 4) – patterns of subgraphs that often recur in complex networks [21]. As different motifs have been associated with different functions, identifying such motifs should allow us to infer their functionality.
2. *Axonal morphology data* allows us to estimate propagation timing across the connections. From propagation timing analysis [23] we can derive properties that optimize salient and timely information processing across a cortical network, and explore how cortical and subcortical areas integrate their operations.

Network motif analysis can characterize connectivities at several different levels (Figure 5). First we will gather cell-to-cell motifs among small subsets of neurons (from a couple to a few, as in Figure 5) at the lowest level of detail (LOD 1, Section 6). Next, we will search for motifs among all the neuron types in Figure 5, at the circuit schematic level (LODs 2-4). To identify the function and properties of each motif, we will compare observed recurring motifs in our data (see Figure 4) with motif-function pairs from various complex networks [1, 2, 15]. After finding motifs and circuit schematics for each distinct small neuronal cluster, we will validate them by comparing them to models of local circuits in the literature [12, 29], from which we expect to infer functions and properties of local circuitry in a complex neural network.

6 Hierarchical analysis

To support structural analysis of the brain network, we will search for modules and their connectivity at five *levels of detail* (LOD):

- LOD 1. *Cell-to-cell connectivity* records the neuron, its type, its dendritic arbor, and the presynaptic neuronal source of each synapse on each of the dendritic segments.
- LOD 2-4. *Cluster-based connectivity* aims to capture local circuitry so as to generalize it over small clusters of source and target neurons and their mutual synapses.
- LOD 5. *Global connectivity of cortical areas/nuclei* [8, 10, 28, 32] is the top level of the hierarchy.

Our approach of pooling locally connected neurons (LOD 2-5) should yield an *architecturally realistic hierarchy of network models*. Figure 5, (adapted from [20, 25]), illustrates cluster-based pooling, as seen in the striate cortex of cat (illustrated), but also in monkey, and human. Here three levels of neuronal pools are identified: wafers (LOD 2), minicolumns (LOD 3), and columns (LOD 4):

- *Wafers* are confined to one of the six layers of the cortex. The cell types, morphologies, and synapses within a wafer are relatively homogeneous.
- *Minicolumns*, ($\sim$30μm dia.), also called *pyramidal cell modules* (PCMs), are commonly hypothesized to be the smallest processing unit of the neocortex [22].
- *Columns* (300-500 μm dia.), within the well-established columnar structure of the cortex [22], can represent pools of 100 minicolumns.
- *Cortical areas*, proceeding up the hierarchy, represent the highest module of the cerebral cortex.

[1] Available at http://www.grub.org.

[2] Available at http://php.indiana.edu/~osporns/graphmeasures.htm.

We will search for higher-level modules (wafers, minicolumns, and columns) and characterize their connectivity in the mouse cerebral cortex. The mouse does not have a striate cortex, and may not exhibit minicolumns, but at the columnar level the barrel fields of its whiskers are prominent and certainly should be picked up by the methods described below. Our search strategy evokes the *inter-wafer connectivity matrix*: In Figure 5, a dendritic arbor within wafer W_i (of PCM I) collects, like an antenna, signals from all of its postsynaptic terminals on incoming fibers (shown as vertically aligned) packed within wafer W_j (in PCM J). Assigning weights to wafer-to-wafer connections defines a coupling matrix $C(I, i|J, j)$ of the direct influence of wafer W_j in PCM J on wafer W_i in PCM I. In electrophysiological terms, the coupling matrix defines the *receptive field* of a wafer.[3] The dendritic arbor or "antenna" collects input from postsynaptic terminals of the neuron, summing postsynaptic currents with appropriate strengths and time delays. Viewed this way as antennae, a small sample of neurons in the wafer can estimate neuronal behavior.

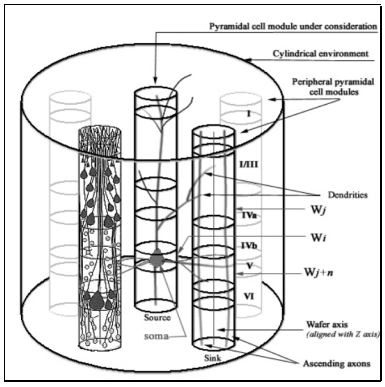

Figure 5: **The modeling environment for a PCM.**

In summary our search for higher-level modules and their connectivity will include:

1. *Wafer-level search.* The first part of our study will identify the invariance properties of a module at a fine grain of clustering, the wafer, under translation and rotation about their vertical axis of symmetry.
2. *Inter-wafer connectivity matrix construction.* Next we will construct and evaluate the degree of translational and rotational invariance of the coupling matrix, to identify minicolumns and columns in homogeneous regions of the mouse cerebral cortex.[4]

7 Seamless integration of biological data from multiple stains and tissues

The great strength of the transgenic mice technology is that it yields promoter-specific staining of unprecedented clarity. But this strength comes at a price: these stain-specific results must be integrated to obtain a fuller picture of the tissue microstructure. Some morphological issues are relatively straightforward, such as the deformation of dendritic trees to reflect the boundaries of the tissue in which they developed. Pyramidal cells, for example, grown beneath a cortical gyral ridge differ from neighboring pyramidal cells grown in sulci. Because we capture a finite-element model of the encapsulating tissue boundaries (Figure 2), we will morph neuronal populations drawn from homologous regions in other mice stained differently to a common finite-element model of mouse brain.

Acknowledgments

Transgenic GAT1-GFP mice, used in this initial study, have been graciously provided by Henry Lester at Caltech. This research was supported in part by the National Science Foundation under grant MRI#0079874 and by the Texas Higher Education Coordinating Board ARP/ATP program under grant #000512-0146-2001.

References

[1] Amaral, L., Scala, A., Barthelemy, M., Stanley, H. Classes of small-world networks. *Proceedings of the National Academy of Sciences, USA*, 97:11149–11152, 2000.

[2] Barabási, A.-L. *Linked*. Perseus Publishing, Cambridge, MA, 2002.

[3] Bower, J. M., Beeman, D. *The Book of GENESIS: Exploring Realistic Neural Models with the GEneral NEural SImulation System*. Telos, Santa Clara, CA, 1998.

[4] Braitenberg, V., Schuz, A. *Cortex: Statistics and Geometry of Neuronal Connectivity*. Springer, Berlin, 2nd edition, 1998.

[5] Chiu, C.-S., Jensen, K., Sokolova, I., Wang, D., Li, M., Deshpande, P., Davidson, N., Mody, I., Quick, M. W., Quake, S., Lester, H. A. Number, density, and surface/cytoplasmic distribution of GABA transporters at presynaptic structures of knock-in mice carrying GABA transporter subtype 1-green fluorescent protein fusions. *J. Neurosci.*, 22:10251–10266, 2002.

[6] Dayan, P., Abbott, L. F. *Theoretical Neuroscience*. MIT Press, Cambridge, MA, 2001.

[7] DeVaul, R., McCormick, B. Neuron developmental modeling and structural representation 1. an introduction to the N++ language, an open stochastic L-system. Technical report, Department of Computer Science, Texas A&M University, College Station, TX, 1996.

[3] Second-order synapses, at a minimum, would be required to model observed electrophysiological behavior, to allow for the interaction of inhibitory and excitatory interneurons.

[4] The coupling matrix strategy should still be applicable where neurons of the minicolumn have migrated.

[8] Fellman, D., Essen, D. V. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1:1–47, 1991.

[9] Hellwig, B. *Dichte und Verteilung prasynaptishcer Boutons. Ein Beitrag zur Synaptologie der Grosshirnrinde*. Thesis, Univ. Tubingen, FRG, 1990.

[10] Hilgetag, C., Burns, G., O'Neill, M., Young, M. Cluster structure of cortical systems in mammalian brains. In Bower, J., editor, *Computational Neuroscience: Trends in Research 1998*. Plenum Press, New York, 1998.

[11] Hines, M. L., Carnevale, N. T. The NEURON simulation environment. *Neural Computation*, 9:1179–1209, 1997.

[12] Kandel, E. R., Schwartz, J. H., Jessell, T. M. *Principles of Neural Science*. Elsevier, New York, third edition, 1991.

[13] Koh, W., McCormick, B. Brain microstructure database system: An exoskeleton to 3d reconstruction and modeling. *Neurocomputing*, 44–46:1099–1105, 2002.

[14] Kozloski, J., F.Hamzei-Sichani, Yuste, R. Stereotyped position of local synaptic targets in neocortex. *Science*, 293:868–872, 2001.

[15] Lee, T. I., Rinaldi, N. J., Robert, F., Odom, D. T., Bar-Joseph, Z., Gerber, G. K., Hannett, N. M., Harbison, C. T., Thompson, C. M., Simon, I., Zeitlinger, J., Jennings, E. G., Murray, H. L., Gordon, D. B., Ben, B., Wyrick, J. J., Tagne, J.-B., Volkert, T. L., Fraenkel, E., Gifford, D. K., Young, R. A. Transcriptional regulatory networks in saccharomyces cerevisiae. *Science*, 298:799–804, 2002.

[16] Lien, J.-M., Morales, M., Amato, N. Neuron PRM: A framework for constructing cortical networks. *Neurocomputing*, 2003. in press.

[17] McCormick, B. Brain tissue scanner enables brain microstructure surveys. *Neurocomputing*, 44–46:1113–1118, 2002.

[18] McCormick, B. Development of the brain tissue scanner. Technical report, Department of Computer Science, Texas A&M University, College Station, TX, March, 2002. http://research.cs.tamu.edu/bnl.

[19] McCormick, B., DeVaul, R. Neuron developmental modeling and structural representation 2. the stochastic model. Technical report, Department of Computer Science, Texas A&M University, College Station, TX, 1996.

[20] McCormick, B., Prusky, G., Tewari, S. Stochastic modeling of the pyramidal cell module. In Bower, J. M., editor, *Computational Neuroscience: Trends in Research 1997*, pages 129–134. Plenum Press, New York, 1997.

[21] Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., Alon, U. Network motifs: Simple building blocks of complex networks. *Science*, 298:824–827, 2002.

[22] Mountcastle, V. *Perceptual Neuroscience: The Cerebral Cortex*. Harvard Univ. Press, Cambridge, MA, 1998.

[23] Nowak, L. G., Bullier, J. The timing of information transfer in the visual system. *Cerebral Cortex*, 12:205–239, 1997.

[24] Paxinos, G., Franklin, K. *The Mouse Brain in Stereotaxic Coordinates*. Academic Press, San Diego, CA, deluxe second edition, 2001. (with CD-ROM).

[25] Peters, V., Yilmez, E. Neural organization in area 17 of cat visual cortex. *Cerebral Cortex*, 3:49–68, 1993.

[26] Pittsburgh Supercomputing Center. http://www.psc.edu/.

[27] Rall, W. Theory of physiological properties of dendrites. *Annals of the New York Academy of Science*, 96:1071–1092, 1962.

[28] Scannell, J. Determining cortical landscapes. *Nature*, 386:452, 1997.

[29] Shepherd, G., editor. *The Synaptic Organization of the Brain*. Oxford University Press, New York, 4th edition, 1998.

[30] Sporns, O., Tononi, G. Classes of network connectivity and dynamics. *Complexity*, 7:28–38, 2002.

[31] Sporns, O. Graph theory methods for the analysis of neural connectivity patterns. In Kötter, R., editor, *Neuroscience Databases: A Practical Guide*. Kluwer Publishers, Boston, MA, 2002.

[32] Young, M., Scannell, J., O'Neill, M., Hilgetag, C., Burns, G., Blackmore, C. Non-metric multidimensional scaling in the analysis of neuroanatomical connection data and the organization of the primate cortical visual system. *Phil-Trans R Soc Lond B*, 348:281–308, 1995.