Visual Data Mining of Brain Cells

Jürgen Symanzik[1], Giorgio A. Ascoli[2,3], Stuart S. Washington[2], Jeffrey L. Krichmar[2]

[1]George Mason University, Center for Computational Statistics 4A7, Fairfax, Virginia 22030, USA.

e–mail: symanzik@galaxy.gmu.edu. WWW: http://www.galaxy.gmu.edu/~symanzik

[2]George Mason University, Krasnow Institute for Advanced Study 2A1, Fairfax, Virginia 22030, USA.

[3]George Mason University, Department of Psychology, Fairfax, Virginia 22030, USA.

e–mail: ascoli@gmu.edu, WWW: http://www.krasnow.gmu.edu/ascoli

Abstract

Little is known about the influence of morphological variablity on the physiological response of brain cells. It is assumed that variables such as the number of branching points in the cell tree–like structure, its area and the change in shape from soma to terminals may influence the neuron behavior differently under different electric stimulation.

In laboratory experiments, it is difficult to separate the electrophysiological effects of biochemical differences (i. e., channel types and their distributions) from those of the morphological differences in a neuron's behavior. Therefore, we have used computer simulations of 3D neuroanatomical data of hippocampal cells while keeping the electrophysiological model constant across different neurons. This allowed us to look specifically at the morphological influence on the neuronal function.

We have obtained neuroanatomical data (particularly, hippocampal pyramidal cells) from public electronic archives. Because of the amount of painstaking work to reconstruct the three–dimemsional structure of a neuron, these archives tend to be very small (ranging from 3 to 50 cells). Additionally, there is the possibility that different archives have significant morphological differences due to the reconstruction technique. Therefore, it is difficult to reach robust conclusions on the relationship between morphological parameters and electrophysiological output.

However, we found that visual data mining techniques combined with computational neural modeling is a very effective way to detect some structure in the data. There are, indeed, morphological variables that can be used to determine whether hippocampal pyramidal cells spike, burst, or have a plateau in response to a given level of input current.

1 Introduction

Despite general agreement among neuroscientists that dendritic morphology plays an important role in shaping cellular physiology, a quantitative analysis of this role has been lacking. Because the task of neurophysiologists, who study the function of neurons by injecting current or voltage into a neuron and measuring the neuron's response, is painstaking and slow, the number of neurons that can be measured in a given preparation is restricted. Often, the electrophysiological responses are recorded, but the neuromorphological data is not mapped. Alternatively, in a neuroanatomical study, the cell architecture is carefully mapped, but the electrophysiological data is not always recorded.

The goal of this paper is to show how visual data mining techniques can be used to explore the neuromorphological effects on the electrical response of neurons. We are applying a computational approach to obtain our data. The procedure is to use three–dimensional neuroanatomical data, take morphological measurements from the data, convert the data into a form that can be used by a computational simulator, and test the physiological response of the simulated neuron. The neuroanatomical data is acquired from public archives. Data consist of rat CA3 hippocampal pyramidal cells (see Figure 1, right, for three such cells) from the Southampton archive (http://www.neuro.soton.ac.uk, Turner et al., 1995). After the simulation, we use XGobi (Swayne et al., 1998) as a visual data mining tool to verify our main hypothesis that neuromorphology shapes neurophysiology. It is expected that differences in physiological response or in firing behavior (i. e., regular spiking, bursting, or plateau potentials) observed within a neuronal family (i. e., CA3 pyramidal cells) can be, in part, attributed to variations in dendritic morphology.

In Section 2 of this paper, we provide a short background on the electrophysiology and the anatomy of hippocampal pyramidal cells. Section 3 describes the simulation studies. We explain our visual data mining approach in Section 4. We end with a discussion and outlook on future work in Section 5.

2 Neurobiological Background

2.1 Hippocampal Pyramidal Cells

Pyramidal cells are the principal cells of the *hippocampus proper*, a portion of the brain present in all mam-

mals. In many species (including rats and humans), the hippocampus has been involved with memory formation. Pyramidal cells have been the focus of a huge number of neuroscientific reports in the last decade. Particularly, CA3 pyramidal cells are part of a highly interconnected network. These neurons receive all their excitatory input through their dendrites, specialized branching structures stemming out of their cell bodies. Pyramidal cells have two types of dendrites, called *basal* and *apical dendrites*, which stem from two opposite parts of the cell body. It is believed that basal and apical dendrites carry different information to the soma. They are also known to have different biochemical and connectivity properties. Typically, a CA3 pyramidal cell has one (or rarely two) apical dendrite(s) and several (2–7) basal trees.

2.2 Electrophysiological Parameters

When a pyramidal cell is excited (for instance by injecting electrical current into its cell body with an electrode), it can react in four different ways (see Figure 1): (1) it remains "silent" (i. e., no response at all); (2) it emits a series of regular, nearly equidistant action potentials ("spikes") (see Figure 1, top); (3) it emits a series of spike trains ("bursts"), i. e., groups of spikes separated by silent periods (see Figure 1, center); (4) it emits a series of "plateau" bursts, where no spike within a train ever goes back to the baseline before the following spike is fired (see Figure 1, bottom). These modes of behavior can each be described by a certain number of parameters. Spiking neurons can be characterized by the frequency of spiking; bursting and plateau neurons can be characterized by their spike frequency within a train, by the inter–train interval, and by the length of the trains (number of spikes per train, or duration). Typically, as more current is injected into a cell, a cell may transition directly from being silent to spiking (these cells are called "spikers", e. g., cell 1 in Figure 2 (a)); or from silent to bursting to spiking ("bursters", e. g., cell 5 in Figure 2 (a)); or from silent to plateau to spiking ("plateau", e. g., cell 8 in Figure 2 (a)). Eventually, at a high enough value of injected current, all cells spike. The same neuron does not usually display both bursting and plateau behavior. Thus, being a spiker, or a burster, or a plateau neuron seems to reflect a *qualitative* firing mode. The electrophysiological behavior of these neurons can be thus further characterized by the miminum amount of current necessary to make the neuron fire in any firing mode, or by the threshold necessary to bring the neuron into spiking mode. Within a certain firing mode, a higher injected current implies a higher firing frequency (see Figures 2 (b), (c), and (d)). The change of quantitative electrophysiological parameters (such as spiking frequency) with change in injected current can itself be considered a quantitative electrophysiological parameter.

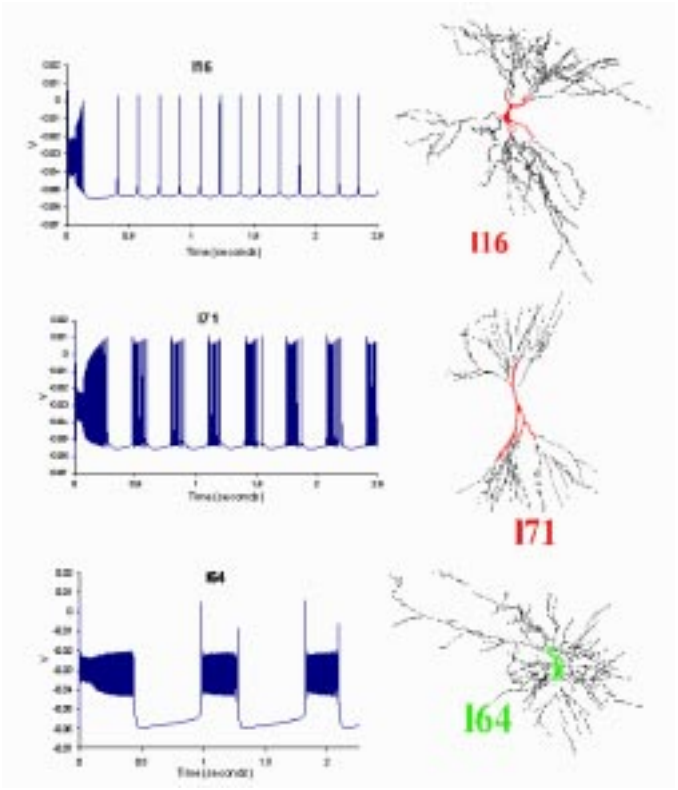

Figure 1: Firing behavior (left) and morphology (right) from cells *l16*, a "spiker", *l71*, a "burster", and *l64*, a "plateau" cell from the Southampton archive.

2.3 Morphological Parameters

A parameter that characterizes neuronal morphology, such as dendritic path length, is defined as a morphometric parameter. We have identified a set of such parameters that can be useful in describing and comparing morphological differences between cell families and within cell families (Ascoli et al., 1997; Ascoli et al., 1998). Examples of these parameters are:

- Average diameter of dendritic branches. This can be measured over the entire dendritic tree or in a specific portion (e. g., within a certain distance from the soma).
- Mean path length. The average distance from a termination point down to the stem of a dendrite. This parameter is correlated with the total dendritic length via the number of stemming trees and the tree asymmetry (Larkman and Mason, 1990).

It should be noted that the list from Ascoli et al. (1997, 1998) only represents a starting point towards

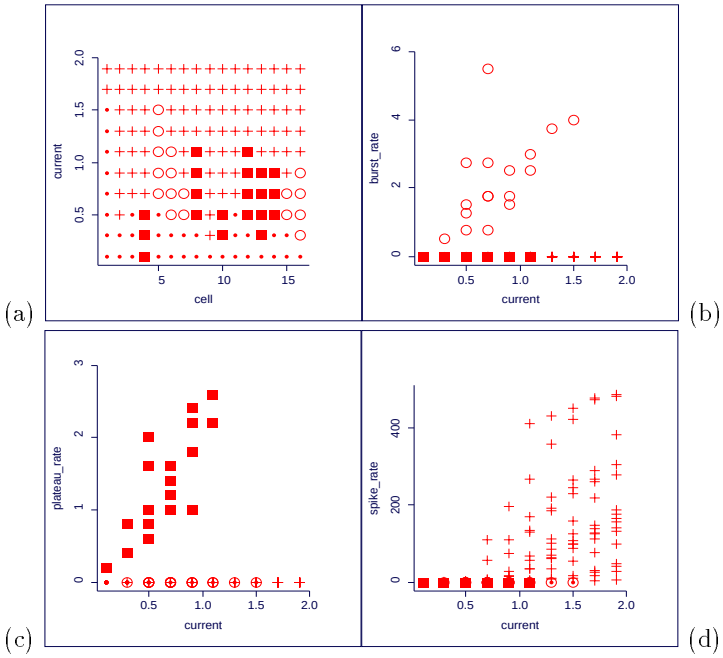

Figure 2: (a) Firing behavior of 16 cells injected with 10 different levels of current in nA. With an increase in current, the burst rate (b), plateau rate (c), and spike rate (d) also increase. Bursts were brushed with a ∘, plateaus were brushed with a ■, and spikes were brushed with a +. A small dot was used when a cell showed no response.

an exhaustive morphometric characterization of neuritic shape. In addition to considering the apical, basal, and both trees for a cell, we also consider morphological parameters in proximity to the soma, i. e., apical, basal, and both trees cut at 50μm, 100μm, 150μm, 200μm, and not cut at all.

3 Simulation Studies

3.1 Known Results

While it is obvious that changes in the concentration and distribution of ionic currents can have a strong influence on neuronal response, it is still an open question as to how much effect morphological differences have on neuronal function and what morphological characteristics are the most influential on neuronal function. In general, smaller cells tend to be more excitable and have higher firing rates. This agrees with our findings (Krichmar et al., 1999; Washington et al., 1999; see also Figures 3 (d) and 4 (left) where cells *l18* and *l60a* that have a large total area spike at a very low rate).

Based on literature reports, morphometric parameters that reflect size, such as dendritic path length and surface area, are expected to correlate sensitively with quantitative electrophysiological parameters (e. g., Larkman et al., 1992; Mainen and Sejnowski, 1996). Bilkey and Schwartzkroin (1990) showed that CA3 pyramidal cells tended to burst as the length of the apical dendrite increased. Our preliminary results concur with many of these previous results; but also indicate that other morphometric parameters in proximity to the soma, such as branching characteristics (e. g., number of branches, branch order for termination points or branch order for segment length) have a significant effect on the neuron's firing mode and this influence needs to be investigated (Krichmar et al., 1999; Washington et al., 1999).

3.2 The Experimental Setting

We started our investigation of how neuromorphology influences neurophysiology by examining 16 CA3 pyramidal cells. Simulations were run using the GENESIS software package (Bower and Beeman, 1994) as previously described (Krichmar et al., 1999). Briefly, neuroanatomical models consist of a list of compartments. Each compartment, which maps to a piece of the neuron's dendritic tree, has a coordinate in 3D space and a diameter. Within each compartment, equations that describe the different ionic currents in the neuron are added. We took these equations from an established model of the CA3 pyramidal cell (Traub et al., 1994). The neuroanatomical files were taken from the Southampton archive and were converted into a GENESIS cell descriptor file. The distribution of membrane properties and active conductances taken from Traub's model were added to the cell descriptor file and adapted to take the realistic morphology into consideration (Krichmar et al., 1999). The correctness of the model was verified by comparing responses of two of the cells (*l51* and *l56a*) to injected somatic current with results from Traub's model and electrophysiological experiments (Hablitz and Johnston, 1981).

In the simulation, cells were injected with current at the soma for five seconds. The level of current ranged from 0.1 nA to 1.9 nA by steps of 0.2 nA. Membrane potential was recorded at the soma. It is apparent that qualitatively different firing modes may be distinguished. Neurons can be spiking (e. g., cell *l16*), bursting (e. g., cell *l71*), or demonstrating a plateau potential (e. g., cell *l64*) (see Figure 1). Quantitative differences were also observed among neurons displaying the same firing mode at a given level of injected current.

The physiological response of each simulated neuron (e. g., spike frequency, interspike interval, bursting vs. non–bursting) to injected current was measured. The effect of morphometric parameters on physiological response was systematically analyzed.

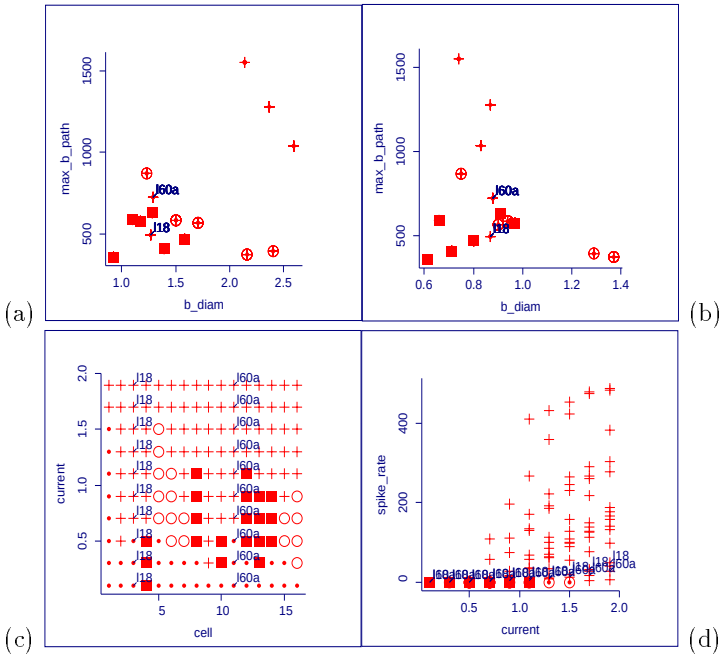

Figure 3: (a) A scatterplot of basal diameter (within a distance of 100μm of the soma) and maximum basal path length (for the entire cell) provides a possible classification for plateaus, bursts, and spikes. (b) When looking at the basal diameter for the entire tree, no such clear classification is possible. Cells *l18* and *l60a* are considered as outliers in (a) and (b). They transition from silent to spiking at 0.7nA (c) and maintain a low spiking rate at higher current levels (d).

4 Visual Data Mining in XGobi

4.1 Introduction to Visual Data Mining

The idea to visually explore data through computer software dates back to PRIM-9 (Picturing, Rotation, Isolation and Masking in up to 9-dimensions; Fisherkeller et al., 1974a; Fisherkeller et al., 1974b), which is the landmark example of early dynamic statistical graphics. Current state of the art technology for visual data mining is available in dynamic statistical graphics packages such as ExplorN (Carr et al., 1997) which is freely available from `ftp://www.galaxy.gmu.edu/pub/software/` and XGobi (Swayne et al., 1998) which is freely available from `http://www.research.att.com/areas/stat/xgobi/`. Main differences between these two packages are that XGobi supports features that are best suitable for relatively small data sets (as in our case) while ExplorN supports features best suitable for larger data sets. We will use XGobi mainly for visual clustering and classification, through a technique called brush–tour strategy that can be successfully used in applications as diverse as human motion data (Vandersluis et al., 1998) and sand particle size data (Wilhelm et al., 1999). The brush–tour strategy refers to an alternation of brushing, grand tour (Asimov, 1985; Buja and Asimov, 1986), brushing, and so on until all interesting results have been visually revealed. Within XGobi, features such as the brush–tour strategy and linked brushing in scatterplots and dotplots have been proven as very successful to detect structure in our morphology data as demonstrated in the next section.

4.2 Visual Data Mining of Pyramidal Cells

When visually exploring the 16 cells using XGobi, we initially started by brushing points that had a different firing behavior with respect to bursts (○), plateaus (■), and spikes (+) (see Figure 2). A small dot was used when a cell showed no response. Since some cells showed different behaviors under different currents, they may be marked, for example, as plateaus for one current and as spikers for another. However, since the underlying morphological parameters are identical, different symbols were overplotted in some of the graphical displays.

When looking at univariate dotplots only, there was no single morphological parameter that resulted in a clear separation of spiker, burster, and plateau cells. However, while cycling through bivariate scatterplots of the morphological paramaters, some structure became apparant. It appears that cells with a relatively small basal diameter (within a distance of 100μm of the soma) and relatively small maximum basal path length (for the entire cell) tend to have plateaus. Cells with medium basal diameter and medium maximum basal path length tend to be bursters. Cells with large basal diameter and large maximum basal path length tend to be spikers (see Figure 3 (a)). While similar results could be observed for other small proximities to the soma, there is no such clear separation when looking at basal diameter for the entire cell and maximum basal path length (for the entire cell) (see Figure 3 (b)).

However, Figure 3 (a) also reveals that there are two outliers, i. e., cells *l18* and *l60a*, that have a small basal diameter and small maximum basal path length but are spikers anyway. When labeling these two cells in the scatterplot of `max_b_path` vs `b_diam` (see Figure 3 (a)), this labeling information is carried over to other linked XGobi plots. The scatterplot of current vs cell (see Figure 3 (c)) reveals that cells *l18* and *l60a* immediately start spiking at a current of 0.7nA while showing no response under smaller currents. Figure 3 (d) reveals that these two cells have spike rates that range among the lowest spike rates among all cells, independently from the injected current.

A natural next data mining step is to investigate whether these two cells have an extreme value in one of their other morphological parameters that could explain

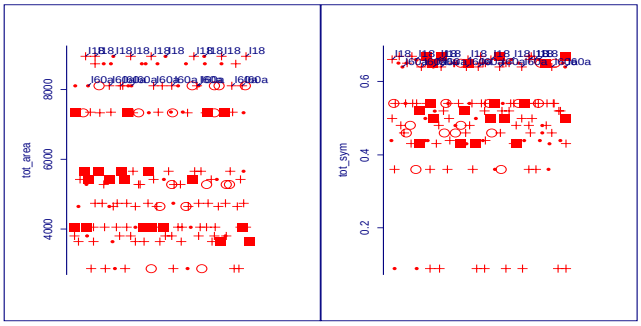

Figure 4: Cells *l18* and *l60a* that are considered as outliers in Figure 3 (a) and (b) have values in the upper range for total area (left) and total symmetry (right) within a distance of 100μm of the soma.

their unusual behavior. And as the dotplots in Figure 4 reveal, cells *l18* and *l60a* have values in the upper range for `tot_area` (total area), `tot_sym` (total symmetry), and total bifurcations (not displayed) within a distance of 100μm of the soma. Obviously, with only 2 cells, it is impossible to state if these high values are pure chance or if one, two, or all three of these parameters (jointly) cause that these cells immediately spike. Clearly, further work is needed.

5 Discussion and Future Work

With only 16 cells $\times$ 10 different currents = 160 observations and 7 morphological parameters $\times$ 5 proximity distances (50μm, 100μm, 150μm, 200μm, and no cut) $\times$ 3 areas (apical dendrite, basal dendrite, and both trees) + 4 parameters for the entire tree = 109 parameters, it seems to be possible to fit a statistical model that describes the data very well but has only little practical meaning. With our visual data mining approach, we believe that we can visually detect the morphological parameters that shape neurophysiology, i. e., a few (interacting) parameters in small proximity to the soma.

After these successful initial steps in visual data mining of hippocampal neurons, the following future work is planned: In a first step, we plan to attempt to reduce and further define the morphometric parameters that shape neurophysioloy. We plan to exploit a large set of real anatomical data using our electrophysical simulations approach. This experiment determines which of the initially explored morphometric parameters, if any, affect the physiology of hippocampal neurons. Based on the previous results, we attempt to develop a statistical model that predicts neuronal function from a given set of morphometric parameters. Finally, we plan to test the predictive, statistical models previously developed on a new set of neuroanatomical data. An independent set of CA3 pyramidal cell data is available from the University of California at Davis archive (`ftp://mossycell.ucdavis.edu/public`).

Acknowledgments

Symanzik's work was supported by a National Science Foundation Group Infrastructure Grant DMS–9631351. Ascoli's work was supported in part by Award No. 00–1 from the Commonwealth of Virginia's Alzheimer and Related Diseases Research Award Fund, administered by the Virginia Center for Aging, Virginia Commonwealth University.

References

Ascoli, G. A., Hunter, L., Krichmar, J. L., Olds, J. L. & Senft, S. L. (1998), 'Computational Neuroanatomy of the Hippocampus', *Society for Neuroscience Abstracts* **24**, Forthcoming.

Ascoli, G. A., Krichmar, J. L., Olds, J. L. & Hunter, L. (1997), 'Computer Generation of Sets of Anatomically Plausible Neurons for Modeling', *Society for Neuroscience Abstracts* **23**(645.5), 1674.

Asimov, D. (1985), 'The Grand Tour: A Tool for Viewing Multidimensional Data', *SIAM Journal on Scientific and Statistical Computing* **6**(1), 128–143.

Bilkey, D. K. & Schwartzkroin, P. A. (1990), 'Variation in Electrophysiology and Morphology of Hippocampal CA3 Pyramidal Cells', *Brain Research* **514**, 77–83.

Bower, J. M. & Beeman, D. (1994), *The Book of GENESIS: Exploring Realistic Neural Models with the GEneral NEural SImulation System*, TELOS/Springer, Santa Clara, CA.

Buja, A. & Asimov, D. (1986), 'Grand Tour Methods: An Outline', *Computing Science and Statistics* **17**, 63–67.

Carr, D. B., Wegman, E. J. & Luo, Q. (1997), ExplorN: Design Considerations Past and Present, Technical Report 137, Center for Computational Statistics, George Mason University, Fairfax, VA.

Fisherkeller, M., Friedman, J. H. & Tukey, J. (1974*a*), PRIM–9: An Interactive Multidimensional Data Display and Analysis System, Technical Report SLAC–PUB–1408, Stanford Linear Accelerator Center, Stanford, CA.

Fisherkeller, M., Friedman, J. H. & Tukey, J. (1974*b*), 'PRIM–9: An Interactive Multidimensional Data Display and Analysis System', ASA Statistical Graphics Video Lending Library (contact: dfs@research.att.com).

Hablitz, J. J. & Johnston, D. (1981), 'Endogenous Nature of Spontaneous Bursting in Hippocampal Pyramidal Neurons', *Cellular and Molecular Neurobiology* **1**(4), 325–334.

Krichmar, J. L., Washington, S. S. & Ascoli, G. A. (1999), 'Effect of Dendritic Morphology on CA3 Pyramidal Cell Electrophysiology', *Neuroscience Letters*. Submitted.

Larkman, A. U., Major, G., Stratford, K. J. & Jack, J. J. (1992), 'Dendritic Morphology of Pyramidal Neurons of the Visual Cortex of the Rat. IV: Electrical Geometry', *Journal of Comparative Neurology* **323**(2), 137–152.

Larkman, A. U. & Mason, A. (1990), 'Correlation between Morphology and Elctrophysiology of Pyramidal Neurons in Slices of Rat Visual Cortex. Establishment of Cell Classes', *Journal of Neuroscience* **10**(5), 1407–1414.

Mainen, Z. F. & Sejnowski, T. (1996), 'Influence of Dendritic Structure on Firing Pattern in Model Neocortical Neurons', *Nature* **382**, 363–366.

Swayne, D. F., Cook, D. & Buja, A. (1998), 'XGobi: Interactive Dynamic Graphics in the X Window System', *Journal of Computational and Graphical Statistics* **7**(1), 113–130.

Traub, R. D., Jeffreys, J. G. R., Miles, R., Whittington, M. A. & Toth, K. (1994), 'A Branching Dendritic Model of a Rodent CA3 Pyramidal Neuron', *The Journal of Physiology* **481.1**, 79–95.

Turner, D. A., Li, X. G., Pyapali, G. K., Ylinen, A. & Buzsaki, G. J. (1995), 'Morphometric and Electrical Properties of Reconstructed Hippocampal CA3 Neurons Recorded in Vivo', *Journal of Comparative Neurology* **356**, 580–594.

Vandersluis, J. P., Cooke, J. D., Ascoli, G. A., Krichmar, J. L., Michaels, G. S., Montgomery, M., Symanzik, J. & Vitucci, B. (1998), 'Exploratory Statistical Graphics for an Initial Motion Control Experiment', *Computing Science and Statistics* **30**, 482–487.

Washington, S. S., Ascoli, G. A. & Krichmar, J. L. (1999), 'Statistical Analyisis of Dendritic Morphology's Effect on CA3 Pyramidal Cell Electrophysiology', *Neurocomputing*. In Press.

Wilhelm, A. F. X., Wegman, E. J. & Symanzik, J. (1999), 'Visual Clustering and Classification: The Oronsay Particle Size Data Set Revisited', *Computational Statistics: Special Issue on Interactive Graphical Data Analysis* **14**(2), 109–146.