

## PROJECT DESCRIPTION

**Introduction to the revision:** This is a revised application based on previous year's submission to the same program which received a "competitive" rating. We addressed the major concerns from the previous round as follows: (1) details provided on image processing (§C.1.3), tracing (§C.2), scaling of validation framework (§C.5.2), and educational activities (§D), (2) included better justification for the use of digital phantoms and included a selective manual validation component (project summary, §A), (3) reduced scope (removed Nissl data and reduced data acquisition extent from 3 to 2 brains per stain) to make the project feasible given 3 years, and (4) graduate students' role more clearly defined (§D). New figures: Fig. 3c&d, Fig. 5, Fig. 6, and Fig. 13.

### A. Introduction

Innovative high-throughput and high-resolution three-dimensional imaging instruments have made the acquisition of unprecedented amounts of sub-cellular level tissue data possible [19, 47, 56]. The *Knife-Edge Scanning Microscope* (KESM), developed and hosted by the PI's research laboratory, is one of the first instruments producing massive three-dimensional data sets [17, 42, 47, 49, 52]. The KESM can section whole small animal organs (e.g., a mouse brain) at sub-micron resolution, generating data at a rate of 180 megabytes/s. There are several mouse brain atlases available to the research community [3, 38, 59, 65], but (1) their resolution perpendicular to the sectioning plane ( $z$ -axis) is limited (50 to 60 times less than that of the KESM), thus fine anatomical structures are hard to disambiguate and trace, and (2) the complexity and amount of data made available poses serious challenges to automated geometric reconstruction, morphological analysis, and dissemination. High-resolution collections are available, however the amount of data that can be imaged is very limited (few neurons from small local regions of the brain) [13, 14, 16, 23, 39].

**Challenge:** The raw three-dimensional data from mouse brain tissue scanned using KESM currently average over 2 TB per brain, posing serious challenges for geometric reconstruction and analysis. These reconstructed models are needed to turn the raw biological data into information and knowledge about brain anatomy and function.

**Key gaps:** Automated algorithms exist for tracing objects of interest (e.g., neurons), but they have limitations in terms of scalability, accuracy, and validation [2, 15, 22, 31, 32, 57]. Most methods are (1) limited to a small number of objects in small data sets, (2) have lower than about 95% accuracy, and/or (3) limited facilities for validation.

**Approach and rationale:** Our approach to address these gaps is to (1) develop a framework to do the tracing automatically while also (2) automating the validation process by certifying the performance using model-based digital phantoms and selective manual validation. Processing large volumes of data requires automated tracing algorithms. Since the automated algorithm may not be 100% correct, validation is needed [61, 76, 78]. However, complete validation often requires that humans perform manual tracing of a volume of data, which defeats the purpose of developing automated tracing algorithms in the first place. Thus, validation based on digital phantoms and selective manual validation is needed.

**Research goals and objectives:** The goal of this project is to develop fast and robust vector tracing algorithms tightly coupled with model-based and selective manual validation. The main objectives of this project are as follows:

1. Develop customized image processing algorithms for serial sectioning imaging,

2. Develop fast, robust tracing algorithms for neuronal and vascular tracing and reconstruction in whole mouse brain KESM data [17, 41, 44, 46].
3. Develop neuronal and vascular morphology models to obtain statistically accurate models of neuronal and vascular networks [34, 35] (international collaboration with Dr. Randal Koene).
4. Analyze and model the sectioning and imaging process of the KESM to generate accurate digital phantoms for model-based validation [26].
5. Develop a large-scale validation framework on (1) manually labeled ground truth from strategically sampled volumes of the data, and on (2) model-based digital phantoms.

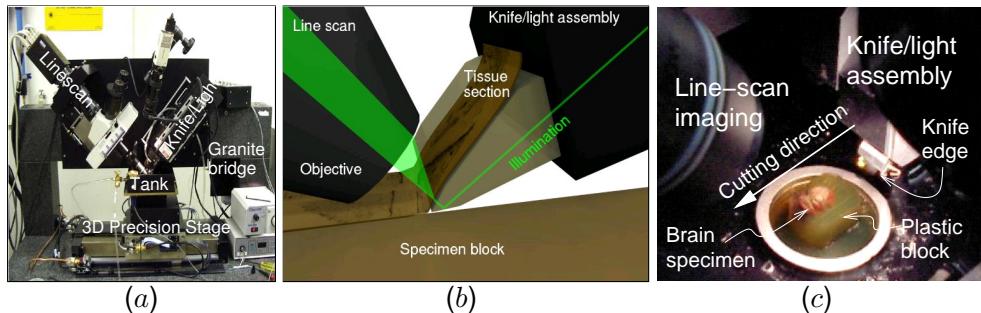
**Intellectual merit, broader impacts, and transformative potential:** Please refer to the project summary.

## B. Background and Prior Work

### B.1. Physical Sectioning Microscopy for Volume Imaging

Critical to our proposed work are massive biological data sets sampled at sub-cellular resolution to which our computation will be applied. The technologies for producing such massive data are actively being developed. Such methods can lead to the complete description of the *connectome*, the study of neural connectivity at the whole-brain scale, identified as a *grand challenge* in neuroscience [67]. Examples of such methods include our Knife-Edge Scanning Microscopy (KESM) [42, 47, 48, 50, 52] (see §B.2), All-Optical Histology [69], Array Tomography [56], Serial-Block-Face Scanning Electron Microscopy (SBF-SEM) [19], and Automatic Tape-Collecting Lathe Ultramicrotome (ATLUM) [28, 29]. Unlike Array Tomography and SBF-SEM, KESM can survey large volumes of biological tissue (whole small animal organs), and KESM is an order of magnitude faster than All-Optical Histology. SBF-SEM has the advantage of ultra high resolution, and Array Tomography is ideal for repeated imaging of a single specimen with multiple immuno stains.

### B.2. Imaging with the Knife-Edge Scanning Microscope



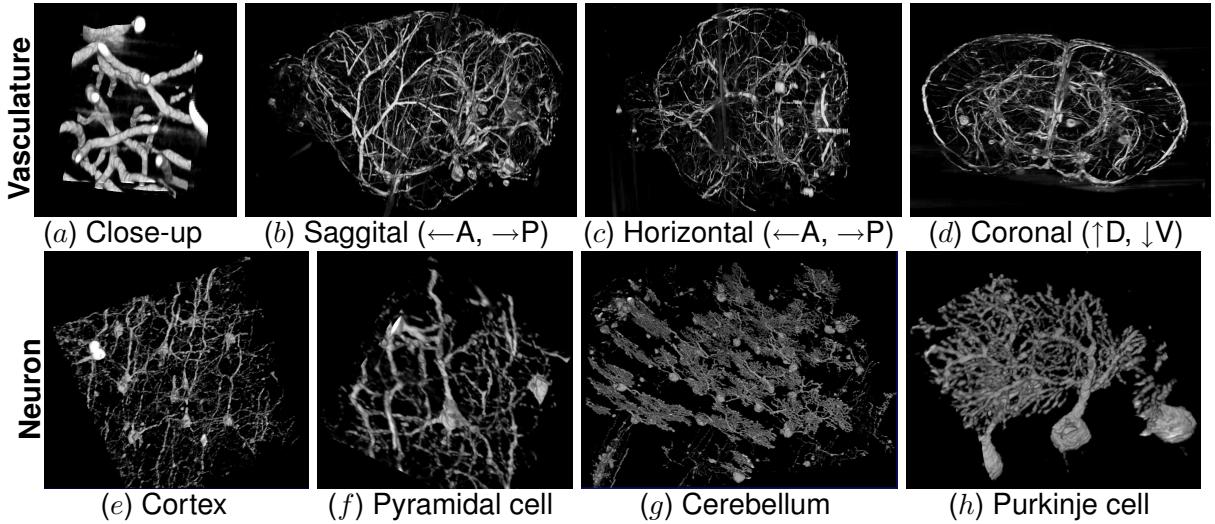
**Figure 1: The Knife-Edge Scanning Microscope.** (a) Photo of the KESM instrument showing line-scan/microscope, knife/light assembly, granite bridge, and 3D precision stage. (b) Specimen undergoing sectioning by knife-edge scanner (thickness of section is not drawn to scale). (c) Close-up photo of the line-scan/microscope assembly and the knife/illumination.

A prototype Knife-Edge Scanning Microscope (KESM, US patent #6,744,572) [47, 48, 52] has been designed at Texas A&M University (TAMU) with support from the National Science Foundation (MRI award #0079874; McCormick, PI), the Texas Higher Education Coordinating Board (ATP award #000512-0146-2001; Keyser, PI), and the National Institute of Neurological Disorders and Stroke (Award #1R01-NS54252; Choe, PI). The instrument, shown in Fig. 1a, is capable of scanning a complete mouse brain ( $\sim 310 \text{ mm}^3$ ) at 300 nm sampling resolution within 100 hours when scanning in full production mode. The instrument comprises four major subsystems: (1) precision positioning stage, (2) microscope/knife assembly, (3) image capture system, and (4) cluster

computer. The specimen, a whole mouse brain, is embedded in a plastic block and mounted atop a three-axis precision positioning stage. A custom diamond knife, rigidly mounted to a massive granite bridge overhanging the three-axis stage, cuts consecutive thin sections from the block. Unlike block face scanning, the KESM concurrently cuts and images (under water) the tissue ribbon as it advances over the leading edge of the diamond knife. A white light source illuminates the rear of the diamond knife, and in turn illuminates the brain tissue at its leading edge, with a strip of intense illumination reflected from the beveled knife-edge as illustrated in Fig. 1b. The microscope objective, aligned perpendicular to the top facet of the knife, images the transmitted light. A high-sensitivity line-scan camera repeatedly samples the newly cut section, imaging a stripe 20  $\mu\text{m}$  wide across the tissue ribbon and just beyond the knife-edge, prior to subsequent deformation of the tissue ribbon after imaging. Finally, the digital video signal is passed through image acquisition boards and stored in a dedicated cluster computing system.

### B.3. Imaging results of neurons and neurovasculature using KESM

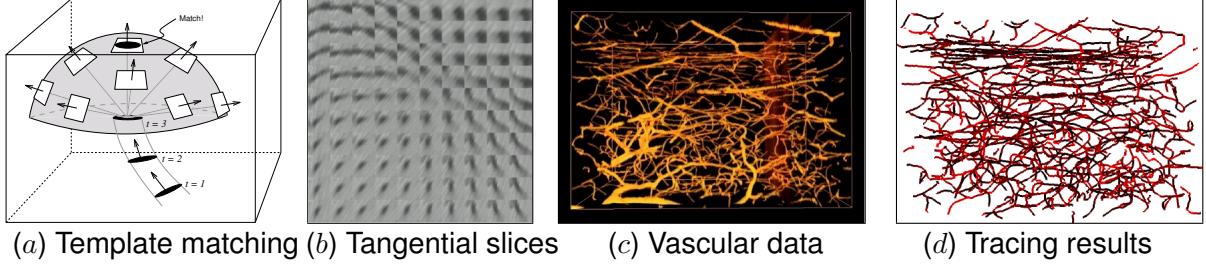
The prototype KESM has been validated on Golgi-, Nissl-, and India-ink-stained mouse brain specimens, and is currently producing high-quality 2D and 3D data. *Nissl staining* dyes the RNA in the cytoplasm of all neurons and the DNA in cell bodies in all cells. However, the dendritic arbors and axons remain unstained. Thus, Nissl staining allows us to reconstruct the distribution of all cell bodies in the mouse brain. Of particular importance is their distribution within the six layers of the cerebral cortex. *Golgi staining*, in contrast, reveals the entire structure of neurons and, as it stains just 1% of the neurons in the tissue, individual neurons can be seen clearly, permitting reconstruction. Finally, *India ink* stains all blood vessels in the tissue sample. Major arteries and veins down to the smallest capillaries are stained, revealing a complex network. Fig. 2 shows volume rendering of the data.



**Figure 2: KESM Data.** Volume visualizations of KESM data stacks are shown for the vascular data set (top row, India ink stain) and the neuronal data set (bottom row, Golgi stain). (a) Close-up of the vascular data. Width  $\sim 100 \mu\text{m}$ . (b-d) Three standard views of the whole mouse brain vasculature (subsampled from high-resolution data). Width  $\sim 10\text{mm}$ . (e-f) Pyramidal cells from the visual cortex. Width  $\sim 300 \mu\text{m}$  (in e) and  $100 \mu\text{m}$  (in f). (g-h) Purkinje cells from the cerebellum. Voxel size is  $0.6 \times 0.7 \times 1.0 \mu\text{m}^3$ . Width  $\sim 500 \mu\text{m}$  (in g) and  $100 \mu\text{m}$  (in h).

### B.4. Rapid tracing of fibrous matter

In order to turn the raw data into a geometric description of the objects of interest (i.e., reconstruction), we are currently developing rapid tracing algorithms for fibrous matter such as neuronal



**Figure 3: Fast Fiber Tracing.** The basic idea behind our fast fiber tracing algorithm is shown. (a) Starting from a seed point ( $t = 1$ ), the direction of the fiber is estimated ( $t = 2$  to  $t = 3$ ). The next point in the fiber trajectory is estimated using a template-based approach. Images are formed as interpolated slices (computed via graphics hardware) through the data volume—sampling in several directions around the current point. (Note that the distance between trace points is exaggerated.) (b) A typical array of slices from the tangential planes are shown. Only a small number of these resemble a circular cross-section of a fiber (upper-right corner). The best matching one is selected as the next point in the fiber trajectory. (c) Volume rendering of cortical vasculatures are shown (width = 350  $\mu\text{m}$ ). (d) Tracing results on (c) are shown. (Pilot results presented in [44].)

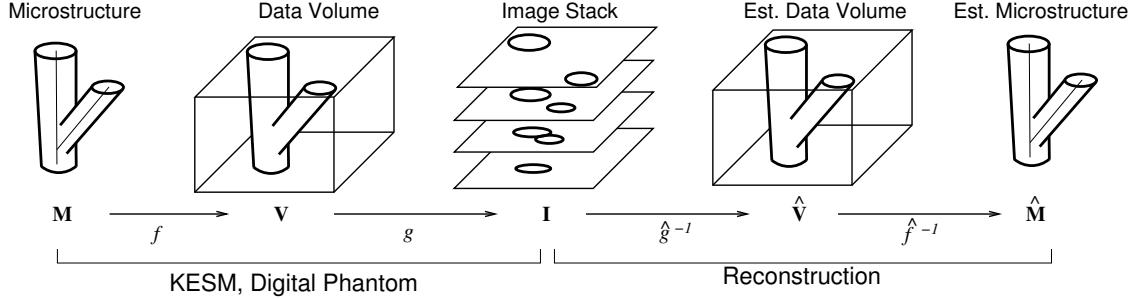
processes and neurovasculature [17, 41, 44, 46]. Fig. 3 shows an overview of our approach, which is broadly classified as a vector tracing method [2, 12]. Our method has been used successfully on KESM and Array Tomography data. See Fig. 3d for preliminary results.

### B.5. Results from prior NSF support

For results from the NSF MRI award (#0079874, PI: McCormick), see §B.2–§B.4 and cited references. The NSF-supported MRI project was further recognized by two subsequent funding from the state and the federal government: the Texas Higher Education Coordinating Board (ATP#000512-0146-2001; Keyser, PI), and the National Institute of Neurological Disorders and Stroke (#1R01-NS54252; Choe, PI). This proposal is distinct from the previous projects, with an added focus on microvasculature (as well as neuronal morphology), and on a neuroinformatics framework focused on automated tracing and model-based validation. Keyser’s NSF supported research (#0220047; Keyser, PI) focused on accurate and robust operations on curved geometry. Efforts from the project resulted in 12 publications (see the CV for details). The PI is currently supported by NSF grant #0905041 (PI: Choe, Co-PIs: Keyser, Abbott), and by grant #0917286 (PI: Keyser). Both projects have just begun on September 1, 2009, so there is no published result yet (2 papers submitted).

## C. Research Plan

We propose to develop automated techniques for extracting and validating structural information from the KESM scans of mouse brain neuronal and vascular networks (Fig. 4). For the automated extraction of structural information, we will improve upon our vector-tracing algorithm for fast and accurate tracing of the objects of interest. Quality control (validation) becomes a critical issue for any automated data processing procedure, and our 3D reconstruction algorithm is no exception [61, 76, 78]. The challenge here is that obtaining full ground truth for validation is not feasible since it could require performing a manual reconstruction of the entire data set. An alternative to manual validation is to use digital phantoms, i.e., synthetically generated data mimicking structures observed in a known ground truth (see [61] for a review). The two approaches have their own limitations, but they are complementary. We will combine the two approaches for accurate,



**Figure 4: Microstructure-to-image Mapping and Reconstruction.** The process by which a microstructure (real or synthetically generated) is turned into a stack of images in KESM and how they are reconstructed is shown. Modeling this process ( $g \circ f$ , composition of  $g$  and  $f$ ) enables the generation of realistic synthetic image data (digital phantoms). On the other hand, the task of recovering the structural descriptions from the image data is basically the inverse:  $\hat{f}^{-1} \circ \hat{g}^{-1}$ , a composition of the segmentation ( $\hat{g}^{-1}$ ) and the 3D reconstruction process ( $\hat{f}^{-1}$ ). (The “ $\hat{\cdot}$ ” symbol indicates that these functions are estimates.) Validation can take three approaches: (1) given  $I$  from KESM and  $\hat{M}_e$  from human labeling, generate reconstruction  $\hat{M}$  from  $I$  and compare it with  $\hat{M}_e$ , (2) given a digital phantom  $I$  with ground truth  $M$ , reconstruct  $\hat{M}$  from  $I$  and compare it with  $M$ , or (3) Given  $I$  and its reconstruction  $\hat{M}$ , use the mapping  $g \circ f$  to create  $\hat{I}$  and show the difference image  $\hat{I} - I$  that can be accepted or rejected.

efficient, high-throughput validation. The research plan is organized as follows: §C.1 will briefly outline the data acquisition and image processing process; §C.2 will give our detailed plan for automated reconstruction; §C.3 will provide details on how to generate statistically accurate and realistic neuronal morphology, to serve as ground truth in generating digital phantoms for automated validation; §C.4 will show how we will model the sectioning and imaging process of the KESM to go from the synthetic ground truth to an actual digital phantom; and §C.5 will explain in detail how validation will be done based on the ground truth.

## C.1. Data acquisition and image processing

We propose to collect additional data from the brain of the laboratory mouse (C57BL/6) beyond the pilot data we have collected (one full Golgi-stained brain and one full India-ink stained brain). This task is now a highly routine and fully automated procedure in our lab. Scanning one brain takes less than 2 weeks, so a total of less than 8 weeks (2 mice per stain, for estimating the across-specimen variability and cross-validation) will be needed to complete this task, while other tasks are carried out in parallel.

### C.1.1. Stain and prepare specimen

We will examine the brain of the laboratory mouse, as it offers certain advantages for study. The brain is small; its cortex is not folded; tracing experiments and imaging techniques can be performed on live mouse brain; the statistical characteristics of its microscopic features have been quantified carefully and extensively [10]; and the laboratory mouse is the leading mammalian model for genomic modeling and gene modification experiments. C57BL/6J mice have been chosen as representatives of the mammalian brain as they are a very common inbred mouse strain used for transgenic studies and aging research. Mice for this study will be obtained from the breeding colony maintained by Co-PI Louise Abbott. Brains from two mice will be used for each stain. Note that part of the brain specimen will be used for pilot runs of the KESM, thus those particular data may not be stored for analysis. Data from two mice for each stain will be collected and analyzed for the production runs. We will use two stains, Golgi and India ink. See Table 1 and [1, 30, 64] for the exact protocols. After staining, the whole brains are dehydrated through a series of graded ethyl alcohols (25%–100%) and then infiltrated with araldite and embedded in blocks [1].

Table 1: Specimen Preparation Protocol.

<b>Golgi:</b> The whole brain is saturated in a Golgi-Cox solution [64], and the cells subsequently dyed by immersion in an ammonium hydroxide solution. One advantage of Golgi staining is that it reveals the <i>entire</i> neuron structure, but only stains a few neurons (1%) in the tissue. Stained cells are randomly selected, so the process is not reproducible across specimens, and thin axons are not stained well. Thus, we will focus on dendritic morphology and variability.
<b>India ink:</b> The mice are perfused transcardially with saline, then 10% formalin and finally with India ink, which fills the cardiovascular system. The ink stays in the blood vessels during processing of the tissue for embedding. We will perform whole body perfusions with fixatives and India ink such that the entire brain is perfused sufficiently.

### C.1.2. Section and image with KESM

Procedures described in §B.2 will be used to generate massive volumetric data sets from each of the stained and embedded specimens described above. The prepared specimens are then scanned by the KESM using a fully automated process. The output of this scanning will be a large volumetric grid of intensity values. Each pass of the knife captures a range of  $x$  values (e.g. 2048 or 4096 pixels), across the entire  $y$  range (9,000-30,000 pixels); typical scans are  $4096 \times 12,000$ . Adjacent slices in  $x$  are taken in a stair-step manner in order to create a full cross-section of the tissue. Thousands of consecutive slices are also taken in the  $z$  direction. The total data produced from a scan of a single mouse brain is currently over 2 TB (using 10X, NA 0.3, Nikon objective).

### C.1.3. Imaging artifact removal and image registration

Knife-edge scanning microscopy has its own sources of noise and artifacts. The primary sources of error include: misalignment of the knife edge (causing a uniform gradient to appear in the image), defects in the knife blade (causing “streaks” in the  $y$ -axis of the image), oscillations in the light source due to AC power fluctuation (causing a regular pattern of light changes along the  $y$ -axis), and knife chatter (causing horizontal lines and loss of most data at those pixels) in the image (see Fig. 5 and §C.4 for details). Image registration issues are minimal due to the way KESM is designed and built.

We have already developed methods for automatically cleaning this data [37, 45] (Fig. 5), retrieving the data-containing region, and aligning the stacks [36]. The full algorithm is as follows:

- (1) Normalize pixel intensity  $n(x, y)$  by row and then by column:  $n(x, y) = \frac{p(x, y)}{\text{median}(\text{row}(y))} \times B$ , where  $p(x, y)$  is the pixel at  $(x, y)$ ,  $\text{median}(\cdot)$  is the median value,  $\text{row}(\cdot)$  is the set of pixels in the row, and  $B$  is the desired background intensity. Column normalization is done in a similar manner. Use of  $B$  allows background illumination to become the same over all images in the image stack.
- (2) Apply step 1 using selective normalization, where pixels are restored to their original value

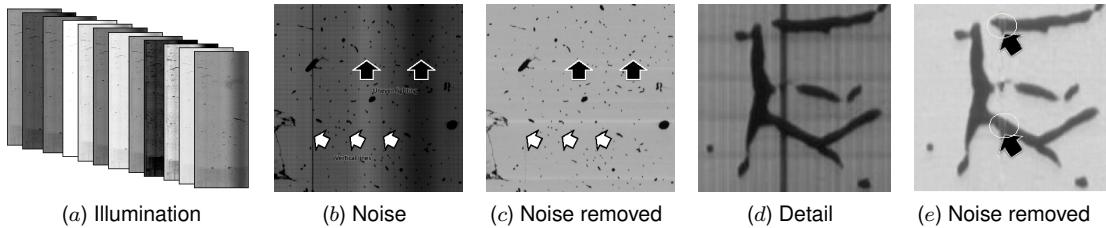


Figure 5: **KESM Image Noise and Noise Removal.** Various forms of KESM image noise and results of noise removal are shown. (a) Uneven illumination across images due the cutting speed variation to avoid mechanical chatter. (b) Local illumination bands (black arrows) and knife-defect bands (white arrows). (c) Noise-removed version of b. (d) Close-up of knife-defect band. (e) Noise-removed version of d using selective normalization.

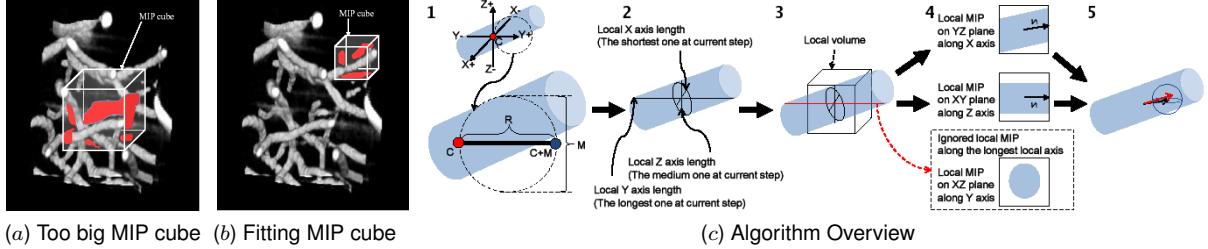


Figure 6: **Vector Tracing with Local Maximum-Intensity Projection (MIP).** (a-b) Local MIP size has to be small enough for unambiguous, occlusion-free projection (a contains occlusions while b does not). (c) The local MIP-based tracing algorithm is outlined. 1. Starting from seedpoint C, longest-axis boundary C+M is found. 2. Boundaries along the other two axes are found. 3. Local MIP cube size is determined. 4. 2D trace is conducted on two of the three MIPs. 5. Results are combined in 3D.

when their normalized values occupy the darker-end peak region (i.e., foreground) in the intensity histogram (Fig. 5d&e). (3) Remove local artifacts due to cutting irregularities using adaptive local scaling:  $p(x, y) = \frac{p(x, y)}{\text{mean}(L(x, y))}$ , where  $L(x, y)$  is a local window around  $(x, y)$ . (4) Crop non-data portion of the image using a boundary detection template to calculate the cut-off column  $x_{\text{cutoff}} = \text{argmin}_x (\sum_{c=0}^w \sum_{r=0}^h |p(x+c, y+r) - p(x, y)|) + \frac{w}{2} - W$ , where  $p(x, y)$  is the pixel,  $(w, h)$  is the template width and height, and  $W$  is the width of the uncropped image. We have found that these algorithms are sufficient in removing most KESM-specific artifacts from the images. Finally, (5) we use stack-by-stack registration using our pattern-matching-based registration algorithm described in [36]. Note that within each stack, images are perfectly aligned due to the way KESM is designed.

## C.2. Develop fast, robust tracing algorithms

Our goal is to extract from the KESM volume data set the individual geometric structures of interest. Fig. 4 shows an overview of the geometric reconstruction process. We have developed a series of vector-based tracing algorithms [27, 41, 43, 54]. Here, we will discuss [27] in detail, while keeping it open to adopt to this project our other approaches as well.

Our maximum intensity projection(MIP)-based algorithm provides a general framework for 3D tracing, and one unique strength is that it can use any 2D tracing algorithm as a subroutine. Thus, the overall power of our MIP-based algorithm can be arbitrarily increased by the use of more powerful 2D tracing algorithms. Furthermore, our algorithm takes only about 30% of the time to trace a unit block compared to a full 3D algorithm. The algorithm consists of four steps, starting from a seed point: (1) boundary detection and local axis length determination, (2) local volume estimation and local MIP processing, (3) 2D tracing of fiber direction, and (4) 3D fiber direction estimation and adjustment. Fig. 6c shows an overview of our approach, and for details see [27].

(1) First, we define  $P_{\text{in}}(\vec{c}, R)$ ,  $P_{\text{out}}(\vec{c}, R)$ , and  $P_{\text{edge}}(\vec{c}, \tau)$  as the probability of a voxel  $R$ -voxel (or  $\tau$ -voxel) apart from trace center  $\vec{c}$  being inside, outside, or on the edge of the fiber, respectively (Gaussian functions are used.) From this, we obtain the edge map  $E(\vec{c}, R, \tau) = \frac{1}{3}(P_{\text{in}}(\vec{c}, R - \tau) + P_{\text{out}}(\vec{c}, R + \tau) + P_{\text{edge}}(\vec{c}, \tau))$ , and  $E(\vec{c}, R) = \max E(\vec{c}, R, \tau)$ . The foreground-background boundary along the three axes  $x$ ,  $y$ , and  $z$  are found by calculating  $\text{argmax}_{R \in [\vec{c}, \dots, \vec{c} + M\vec{e}]} E(\vec{c}, R)$  where  $M$  is the maximum fiber width and  $\vec{e}$  is the unit vector along one of the principal axes. (2) From these boundaries, the local MIP volume is created, and MIPs generated on the XY, YZ, and XZ planes. Among these, the projection along the longest fiber direction is ignored since it does not contain fiber direction information. (3) Tracing in 2D is done using a multiscale filter using a Hessian matrix  $H_\sigma(x) = \begin{bmatrix} I_{xx}(x) & I_{xy}(x) \\ I_{yx}(x) & I_{yy}(x) \end{bmatrix}$ , where  $I_{uv}(x) = \sigma^2 \frac{\partial^2 G(x, \sigma)}{\partial u \partial v} I(x)$ , where  $\sigma$  is the scale, and  $G(x, \sigma)$  is a Gaussian with standard deviation  $\sigma$ , and  $I(x)$  is the voxel  $x$  in the data volume

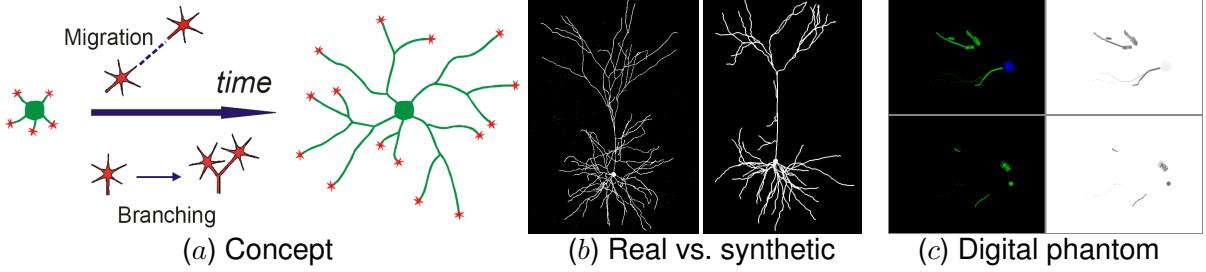


Figure 7: **NETMORPH Framework for Simulating Neuronal Growth.** An overview of the approach and pilot results are shown. (a) An overview of how model neurons are grown is shown. (b) A comparison of a real neuron (left, from Svoboda’s lab) and a synthetic neuron generated from NETMORPH (right) are shown. (c) Preliminary results from the SLICE module is shown. The SLICE module generates digital phantoms (currently, no noise model is included).

I. The eigenvector  $\vec{v}_1$  corresponding to the smallest eigenvalue  $\lambda_1$  indicates the fiber direction. To elaborate, first a fiber-likeness function  $V(x, \sigma)$  is defined as 0 when 2nd eigenvalue  $\lambda_2 > 0$ , and as  $\exp\left(-\frac{\lambda_r^2}{2\beta^2}\right)\left(1 - \exp\left(-\frac{S^2}{2c^2}\right)\right)$  otherwise, where  $\lambda_r = \frac{\lambda_1}{\lambda_2}$ , and  $S = \sqrt{\sum_{i=1}^2 \lambda_i^2}$ , and  $\beta$  and  $c$  control the sensitivity of the Hessian filter (0.5 and 0.25, in our experiments). The maximum value along  $V(x, \sigma)$  at scale  $\sigma$  gives the eigenvector  $\vec{v}_1$ , which is the fiber direction (see [24] for details). (4) Finally, from the 2D tracing on the two MIP planes, the 3D fiber direction is estimated as follows. Assume that XY and YZ were the two chosen planes. Let  $\vec{v}_{xy} = (x_1, y_1, 0)$  and  $\vec{v}_{yz} = (0, y_2, z_2)$  be the eigenvectors of the two respective MIP planes. The 3D fiber direction is derived as  $\vec{v}_{xyz} = (x_1, \frac{y_1+y_2}{2}, z_2)$ . From this, the next trace center  $\vec{c}'$  can be calculated as  $\vec{c}' = \frac{\vec{v}_{xyz}}{\|\vec{v}_{xyz}\|}\rho + \vec{c}$ , where  $\rho$  is the step size (currently set to 3). The last step corrections for discretization error, where we use a momentum operator to calculate the adjusted center  $(\bar{x}, \bar{y}, \bar{z}) = \left(\frac{M_{100}}{M_{000}}, \frac{M_{010}}{M_{000}}, \frac{M_{001}}{M_{000}}\right)$ , where  $M_{pqr} = \int_z \int_y \int_x x^p y^q z^r I(x, y, z) dx dy dz$  (see [9] for details).

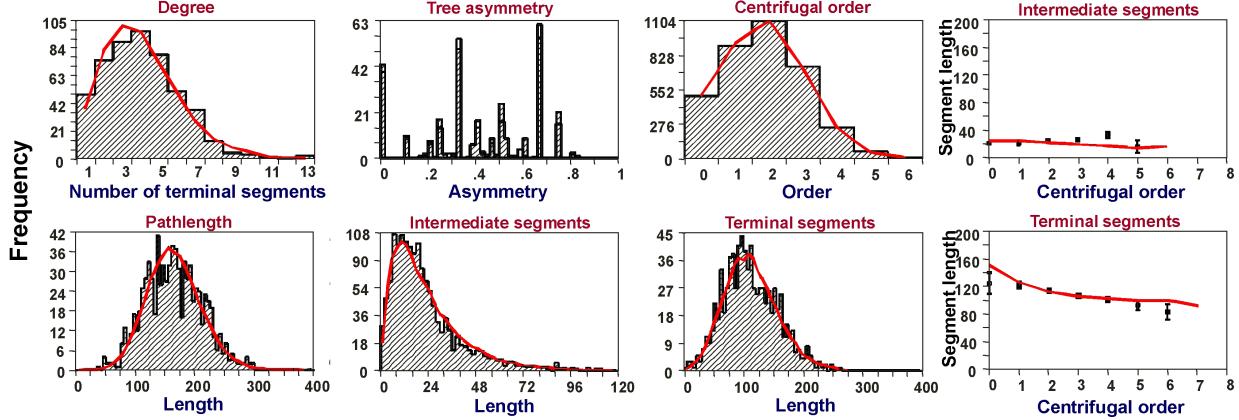
Our algorithm, due to the use of a 2D Hessian matrix, has computational complexity of  $O(n^2)$ , while full 3D tracing algorithms based on 3D Hessian matrices have  $O(n^3)$  complexity, where  $n$  is the scale of the Gaussian filter. In practice, our algorithm takes only about 30% of the time of full 3D versions. Initial implementation of this and related approaches has proven successful [27, 41, 43, 54], and we will modify and improve it to trace a wider variety of fibrous data.

### C.3. Develop neuronal and vascular morphology models

In order to generate realistic digital phantoms, we need to start with realistic ground truth, i.e., synthetic neuronal and vascular networks that are morphologically accurate. NETMORPH, developed by Randal Koene (collaborator on this proposal), provides an excellent platform [34, 35].

**Neurogenesis with NETMORPH:** NETMORPH is a framework for the large-scale simulation of developing neuron morphology and neuronal network connectivity in 3D (Fig. 7). NETMORPH is unique in that it incorporates development explicitly, using stochastic models that are grounded in experimental observation [71, 72, 73, 74, 75]. Neuronal morphogenesis is simulated from the perspective of the individual growth cone. The various actions of the growth cone, such as elongation, branching and turning, are described in a stochastic manner that also supports a phenomenological implementation of biophysical processes involved in neurite outgrowth, such as competition for resources between different growth cones of a dendrite or an axon. The outgrowing neurites can interact with their simulated physical environment.

In this way, neurons with realistic axonal and dendritic morphologies, including neurite curvature, are generated. Characteristic neuron morphologies, such as those of basket cells and pyramidal cells (see Fig. 7b) that clearly resemble observed real morphologies emerge when the



**Figure 8: NETMORPH Results.** Various statistical distributions of morphological features from the NETMORPH-generated neurons are shown (segment length, branching order, number of terminal segments, etc.), where a close match is shown not only in the summary statistic (mean and variance) but also in the actual distribution.

parameters of development models are fine-tuned with the aid of a genetic algorithm. Simulated development with NETMORPH includes synaptogenesis, so that networks are formed.

The few modeling frameworks that exist for the generation of neuronal networks either do not consider detailed neuronal morphology [11, 66, 70] or, if they do, do not take into account the developmental aspects of neuronal morphogenesis and synapse formation [5, 6, 25].

Fig. 8 shows preliminary results from NETMORPH, where close correspondence between the statistical distribution of features (such as segment length, branch order, etc.) in real neurons and their synthetic counterparts are highlighted. Properties of the synthetic neurons generated by NETMORPH match with those of the real neurons in terms of the full probability density function, not just summary statistics such as mean and variance.

We will extend the NETMORPH framework to support multiple cell types and their networks in major regions such as the visual cortex, cerebellum, and hippocampus.

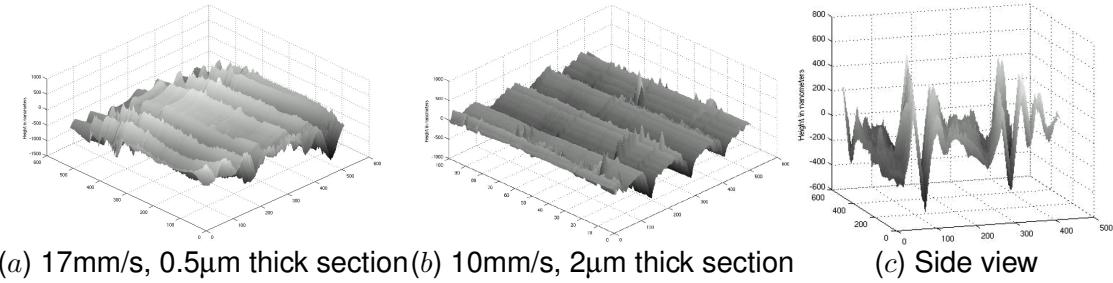
**Vasculogenesis with extended NETMORPH:** We will extend the NETMORPH framework to allow growth of vascular networks. The same computational framework in NETMORPH will be used, except for the specific growth rules which we will adapt from the vasculogenesis and angiogenesis literature [4, 60, 68].

The growth of the vascular network is in many ways similar to that of the neurons. Endothelial cells migrate, leading the way of capillary sprout growth, which could be similar to migration of neurons; and the capillaries branch and reconnect through anastomosis and mitosis [4]. Specifically, we will consider diffusion of angiogenetic factors, endothelial cells' uptake of these factors, and stochastic sprouting and sprout direction determination [68].

Once the realistic neuronal and vascular networks are generated, the SLICE module will be used to generate digital phantoms. See §C.4 for details on modeling the sectioning and imaging process, to provide accurate functions and parameters for the SLICE module.

#### C.4. Analyze and model the sectioning and imaging process of the KESM

As mentioned above, for large-scale automated validation, the use of ground truth produced by manual labeling is impossible. In addition, the model-matching procedure may prove time-



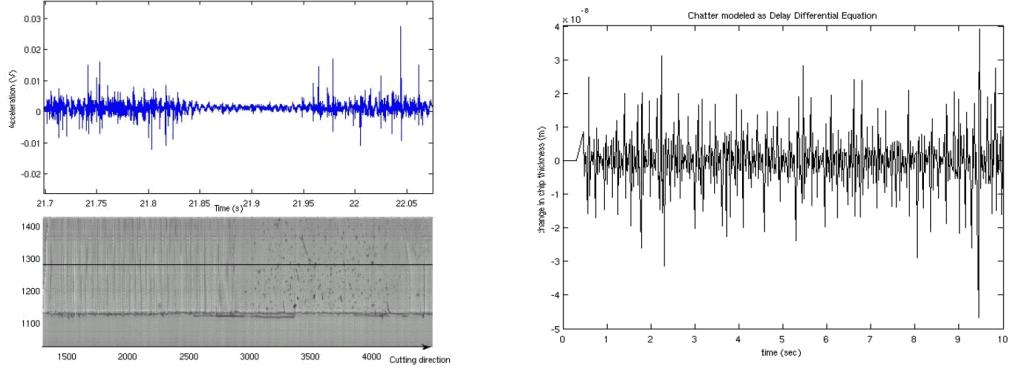
**Figure 9: Surface Profiling with Atomic Force Microscopy.** Various surface profile measurements using an atomic force microscopy (AFM) are shown. Different cutting velocities and section thickness have been tested (a and b). (c) shows a side view of a typical scan. Note that the surface features show a regular peak and valley (chatter), with the height of the peak on the order of several hundred micrometers. Pilot result from [26].

consuming due to the extensive graph-matching component in the algorithm. Thus, even if ground truth existed for the entire data set, validation can take an exorbitant time. One way to circumvent this problem is to use digital phantoms [18, 40, 61].

**Estimate the sectioning and imaging process:** Digital phantoms are synthetically generated, realistic data, produced from a known ground truth model (Fig. 4). Given a known ground truth  $M$ , the 3D volume  $V$  is generated in a straight-forward manner, and then image formation, distortion, and noise models are used to produce the final product, the image stack  $I$ . The process simulates physical data acquisition in instruments like the KESM. Thus, we need to estimate the volume production function  $f(\cdot)$  and the imaging function  $g(\cdot)$ . The estimation can be based on KESM data  $I$  and manually labeled ground truth  $\hat{M}_e$ . The advantage of this approach is that digital phantom generation is a straight-forward causal process, thus modeling it is easier than the inverse problem (reconstruction), which may need additional constraints as the inverse problem is ill-posed (see e.g., [62]). The main research issue involves estimating the imaging process  $g(\cdot)$  (the volume generation process  $f(\cdot)$  is relatively easy). This process involves optics, luminance irregularities, tissue sectioning artifacts (variation in thickness, chatter in the sectioning process, and minor tissue distortion), and noise. The main source of noise is caused by the mechanics of the sectioning process, called “chatter”. Due to chatter, variation in the thickness of the sectioned tissue results, leading to noise in the image.

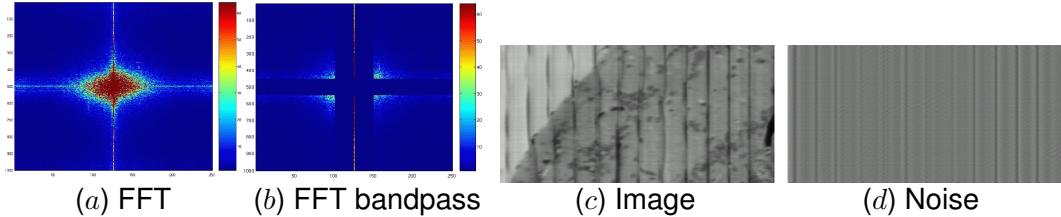
**Collect data from the sectioning and imaging process:** We have already gathered preliminary data to characterize the mechanical properties of chatter, such as (1) surface profile measured with Atomic Force Microscopy (Fig. 9), (2) force estimated with an accelerometer as compared to the actual imaged tissue section (Fig. 10a), and (3) elasticity of embedding material and embedded tissue using a durometer [26]. It is also important to quantify the optical artifacts, and our initial results using the Fast Fourier Transform are promising (Fig. 11). Due to chatter, dark bands appear at a regular intervals (Fig. 11b, left), which can be extracted by bandpass filtering in the Fourier domain (Fig. 11a, right, and results shown in Fig. 11b, right).

**Construct a nonlinear dynamic model of the sectioning process:** Using the above data, we will construct parametric models for each of these processes and use optimization methods to fit the parameters, with which we will generate highly realistic digital phantoms. We have begun modeling the sectioning process, with promising initial results [26, 51] (Fig. 12, Fig. 10b). These models are largely based on nonlinear dynamical models of the sectioning process studied in the machining literature [33, 77]. Orthogonal cutting (sectioning), as shown in Fig. 12, is one of the simplest cutting process. In this model, various forces and other factors such as cutting speed



(a) Accelerometer reading and tissue image      (b) Surface profile from model

**Figure 10: Measuring and Modeling Chatter.** (a) Piezo-electric accelerometer reading (top) and the corresponding image acquired from the cutting session (right) are shown. The part with higher acceleration is followed by a lower amplitude region in the middle, again followed by higher amplitude in the acceleration. The lower amplitude region corresponds to the part where actual brain tissue is embedded, while the other regions are plastic only. Pilot result from [26].



**Figure 11: FFT Analysis of Repeating Noise Signature.** (a) 2D Fast Fourier Transform (FFT) of images obtained from the KESM is shown, along with (b) the filtered version to extract chatter features. (c) The original image used in calculating the FFT in (a) is shown with (d) the extracted chatter features. The dark repeating horizontal bands can be extracted with this approach. Pilot result from [26].

and section thickness come into play. The cutting tool exerts force  $R$ , and the work piece gives resistance  $R'$  (Fig. 12a). The force  $R$  can be decomposed into the cutting force  $F_c$  and thrust force  $F_t$ ; and the resistance  $R'$  into the shearing force  $F_s$  and the friction force  $F_f$  (Fig. 12b). The various angles such as the rake angle  $\alpha$ , shear angle  $\phi$ , and the friction angle  $\tau$  all play an important role in characterizing the cutting process. We will extend a general oscillator equation to model the variation in the knife position  $x$  due to chatter:

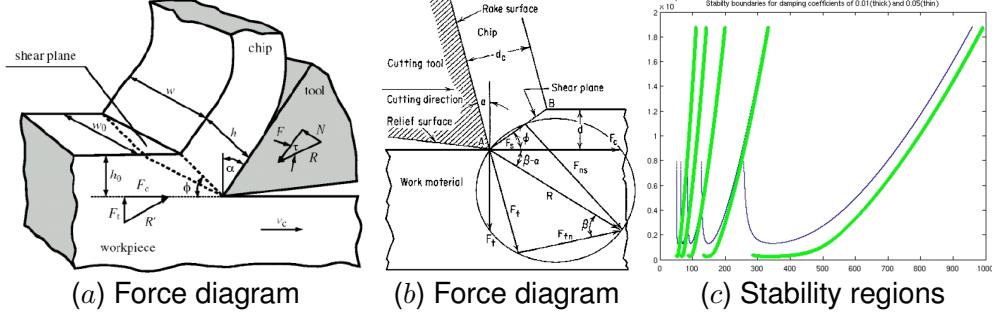
$$\ddot{x} + 2\xi w_n \dot{x} + x = k(x(t) - x(t - \tau)),$$

where  $\xi$  is the chip thickness ratio,  $w_n$  is the chip width,  $k$  is a proportionality constant, and  $\tau$  is the distance between the position of the knife and the position of the surface of the chip at that location. Fig. 10b shows pilot results from this model, which closely resembles the accelerometer reading in Fig. 10a.

### C.5. Develop a large-scale validation framework

We will develop a large-scale validation framework where a recursive, iterative improvement of the reconstruction algorithm and the validation procedure are interlocked together.

We will develop our validation approach in two stages. In the first, we will use small volumes of data, performing validation over these volumes. Then in the second stage, we will scale up the validation to large-scale volumes. Both stages will use an iterative improvement process to improve



**Figure 12: Various Forces Involved in the Cutting Process.** The various forces involved in the orthogonal cutting process are shown in (a) and (b), from [77] and [33], respectively. With a dynamic nonlinear model of cutting, we can determine stable boundaries of the cutting process as a function of  $M$  (a variable inversely proportional to the cutting speed,  $x$  axis) and  $K$  (cutting coefficient,  $y$  axis), for different damping coefficients (thick green curve = 0.01, thin blue curve = 0.05). Adapted from pilot results in [26].

our reconstruction approach. The reconstruction algorithm and the digital phantom function will be co-learned using an Expectation-Maximization framework [53], through a recursive improvement process. We will repeatedly refine our reconstruction and digital phantom algorithms to improve on the given metrics.

### C.5.1. Selective small-scale validation

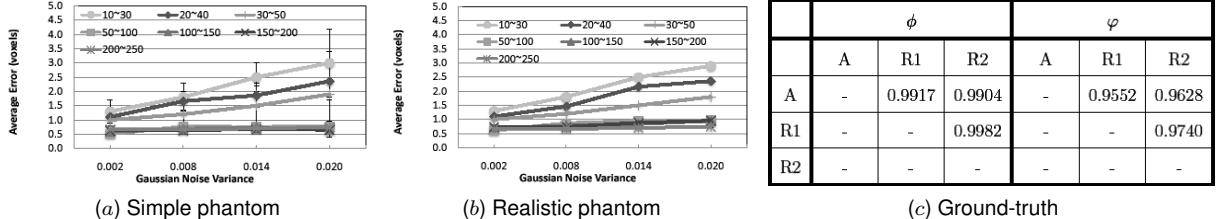
As a first step, we will operate on small volumes of data, for example, a  $300\mu\text{m} \times 300\mu\text{m} \times 300\mu\text{m}$  cube of mouse brain. Consider Fig. 4 which illustrates the image acquisition and reconstruction process. Given the data volume (image stack  $I$ ) obtained from KESM, we will have domain experts (from Co-PI Abbott’s lab) label the data to provide the ground truth  $\hat{M}_e$  (note that the domain expert’s morphological labeling is in itself an estimation of the actual ground truth  $M$ ).

Concurrently with manual labeling, the data volume  $I$  will be put through our segmentation and reconstruction algorithms, producing the estimated microstructure  $\hat{M}$ . Validation can now be done by comparing  $\hat{M}_e$  (by the domain expert) and  $\hat{M}$  (by the reconstruction algorithm). We will use mutual information for the comparison:  $I(\hat{M}_e, \hat{M}) = H(\hat{M}_e) + H(\hat{M}) - H(\hat{M}_e, \hat{M})$ , where  $H(\cdot)$  is the Shannon entropy  $-\sum_{p \in P} p \log p$ , and  $H(\cdot, \cdot)$  the joint entropy. We will also employ measures such as normalized mutual information, entropy correlation coefficient, and gradient [63].

The key research issue in this task is how to map corresponding parts in the geometric microstructure model of the ground truth  $\hat{M}_e$  and that of the estimate  $\hat{M}$  in order to estimate the joint probability. We will employ two approaches, here.

First, we will use an approach for **1. Morphology comparison**. For this, we will consider the geometric microstructure models as graphs of vertices and edges, and employ approximate graph matching (often called “attributed graph matching”) algorithms (e.g., [21, 55]) to align the models being compared. Then, we will treat the morphometry (diameter, in-degree and out-degree, volume, length, etc.) associated with the pairs of vertices and edges as joint events, from which we can calculate the probability distributions. **2. Imaging comparison.** Given reconstruction  $\hat{M}$  from KESM images  $I$ , calculate  $g \circ f(\hat{M})$  to regenerate an estimate of the image stack. This reconstructed volume will be directly compared with the digital phantom of manually labeled ground truth ( $= g \circ f(\hat{M}_e)$ ). In addition, we can compare our reconstruction directly with the original KESM stack of images. For these comparisons, we intend to use one or more of the numerous existing image comparison algorithms: select a good one, rather than develop yet another one on our own.

This second approach has three advantages. First, by employing more standard image-based comparisons, the definition and computation of metrics for comparison is substantially easier.



**Figure 13: Pilot Validation Results.** Systematic validation results with varying image contrast and noise level for digital phantoms made of (a) simple, geometric structures and (b) realistic vascular trees are shown. (c) Validation result A against two manual tracing R1 and R2 on real vascular data are shown ( $\phi$  means length difference, and  $\varphi$  means centerline deviation. Values in the table are the correlation coefficients. Note that the results are very similar between a and b that use different types of phantoms, and tracing is close to perfect as can be seen in c. (Pilot results presented in [27].)

Second, by comparing reconstructions based on stacks, we eliminate the effects of structural variation that occurs at either too high of a frequency or too fine of a precision to actually be contained in the image scan. That is, the method is robust to variations between  $\hat{M}$  and  $\hat{M}_e$  that cannot be determined from the scanned data. Third, the image-based comparison allows for comparison of the reconstructed data directly to the raw image scans, allowing the iterative improvement process to progress farther than might be possible with synthetic data, and removing the limitation that all data be hand-labeled for verification.

On the other hand, this approach has disadvantages in that finer structures that are known to exist but cannot be known from the raw data will be ignored, even if expert knowledge determines they should be there. That is, there may be some artificial limitation in the amount of detail an expert may provide to the process. Also, there is a heavy reliance on the accuracy of  $g(\cdot)$ , which means that differences could be due to problems other than in  $\hat{M}$ .

### C.5.2. Conduct large-scale validation using model-based synthetic data

There is a need to scale the validation process to larger volumes of data. Several of the structures we are dealing with span very large extents in the volumetric data set (beyond the level that can be kept in memory at one time), and we must be able to reconstruct such structures. At the same time, the size of data we are dealing with makes manual reconstruction far too time-consuming. A different mode of validation is necessary.

We propose to address this scaling issue through two approaches. First, we can expand on the imaging comparison (see §C.5.1), by comparing directly to the KESM stack. The process should be the same as for smaller data sets, though it is likely that the image comparison metrics might need to change to account for variations across larger scales. Second, we will make use of synthetic data sets, which we discuss briefly below. The synthetic data sets will allow us to describe features of the reconstructed geometry at a higher level of abstraction, and ensure that we are reconstructing the overall features of microstructures, and not just the low-level geometry.

With small but accurate manually labeled ground truth reconstructions and a reliable digital phantom generator in hand, we can conduct a large-scale validation with model-based stochastically generated data volume. The intention here is (as is often done on a smaller scale) to generate synthetic data for which we know “true” results, and confirm whether our reconstruction accurately determines the structure. The morphological characteristics of neurons are governed by branching degree, length, diameter, and surface area of segments constituting the dendritic and axonal arbor [7, 8]. However, these latter are close to raw data, thus a more structured approach may be necessary to describe the “gist” of the structures that we wish to represent (this is discussed more below). Rule-based approaches for the stochastic generation of synthetic morphologies are

based on these raw metrics [7, 8, 20], and thus may not be able to replicate higher-level structures found in biological microstructure.

Across larger regions, the key features of a neuron or of vasculature are not the fine-scale geometric details, but rather overall measures such as extent, shape, proximity, etc. For example, a reconstruction of a Purkinje cell should accurately model the flattened nature of the dendritic arbor and the nearly parallel “stacking” of sets of cells. Such larger features are especially true in terms of usefulness for much simulation-based research, where the goal is to understand how sets of cells interrelate, rather than to necessarily reconstruct every cell. By controlling the synthetic generation, we can ensure that the reconstruction accurately reflects the overall characteristics of the cells being reconstructed. See Fig. 13 for preliminary results with digital phantoms and selective manual validation.

## D. Broader Impact Activities

**Student training and under-represented groups:** This project will support the training of three Ph.D. students (two directly supported, and one other supported by other funds). Our team’s previous work in this area also has a strong history of involving Master’s-level students (not directly funded; seven M.S. theses have come from the group) and undergraduate students (PIs Keyser and Choe have supervised thirteen undergraduates in the past 7 years) under Research Experience for Undergraduates (REU) program (REU site, #0353957; PI: V. E. Taylor), and the Computing Research Association(CRA)’s Distributed Research Experiences for Undergraduates (DREU) that targets under-represented groups in CS. We will continue our effort in this direction.

The resulting database and geometric reconstructions also gives us an opportunity to reach out to K-12 students. PIs Keyser and Choe have had significant experience in interactions with high-school level groups, with Keyser having given multiple presentations to such groups, and Keyser and Choe organizing a computer science contest/outreach day (hosted by the department) for high school students each of the last three years. We will organize a computational neuroscience contest based on simplified versions of our database and tools.

Our past efforts have demonstrated that we are able to incorporate under-represented groups in the project, such as women (currently two of our graduate students are female, and two women graduated recently) and minority students (primarily at the undergraduate level). We will continue to involve and reach out to these groups in our proposed project.

**Dissemination:** A separate NSF-funded project (#0905041: *CRCNS data sharing: Whole Mouse Brain Neuronal Morphology and Neurovasculature Browser*) is providing a method for distribution, browsing, and basic visualization of whole data sets. In this proposed project, we will add on to that system the geometric reconstructions and validation data (in years 2-3) we develop here. As the reconstruction code becomes mature, we will release it in an open-source environment, such as <http://sourceforge.net>. Software tools will be continually maintained through an active project on SourceForge, which provides support for development activity such as bug tracking and user forums. Existing code (for noise removal) from our group has already been released on SourceForge. We will actively advertise this new resource through various channels: mailing lists, news briefs in scientific publications (such as Science, Nature, and also popular science magazines), personal contact, and demos and exhibitions at scientific meetings. We will also organize short courses and workshops to expand and support the user community. For the above, co-PI Choe will depend on his extensive experience in most of the above activities (Topographica software project <http://topographica.org>, book-related web page and promotion <http://computationalmaps.org> [58], exhibitions as part of the NIH/NIMH Human Brain Project, and

a minisymposium at the Society for Neuroscience meetings [we organized one in 2008], etc.).

## E. Project Management

**Management and Task breakdown:** Dr. Randal A. Koene, Director, Neural Engineering Corporation (also Director, Department of Neuroengineering, Fatronik Foundation, San Sebastian, Spain), will work with the Texas A&M team, to provide his expertise on the NETMORPH platform. The PI, Co-PIs, and the collaborator will manage the project as follows: Keyser (PI: §C.1.3, §C.2, §C.5), Choe (Co-PI: §C.1.2-3, §C.3, §C.4), Abbott (Co-PI: §C.1.1, §C.5.1, and providing continual feedback from a biologist's point of view), and Koene (§C.3). In addition to the two supported graduate students, we anticipate having at least one student work on this project who is supported via either fellowship or teaching assistantship. The graduate students will, roughly, focus on the scanning and cutting process (§C.1, §C.4), the tracing and reconstruction (§C.2), and the morphology modeling (§C.3 - working closely with Koene), with all three involved with validation (§C.5).

**Sustainability:** In order to ensure continued support of the informatics platform developed through this project, we will design and implement a protocol for software development, documentation, and education. All members of the development team (both internal and external) will be trained to abide by this protocol. Involvement of external developers as well as department-funded graduate students will ensure continued support beyond the funding period. We expect maintenance to require less effort once the initial platform is implemented during the project period, and we believe we can build on the work of our recently funded CRCNS data sharing project to make dissemination easier. We will also seek continued funding from follow-up grants and from collaborations.

**Target audience:** The target audience of our project is diverse. Stephen J. Smith (Stanford, Array Tomography), George Karniadakis (Brown University, vascular transport modeling), and Graziano Fiorito (Stazione Zoologica, Italy, octopus neuronal morphology) will be the immediate users of our framework. Other groups who have shown interest in our work include Peter Hunter (CellML, Univ. of Auckland, New Zealand), Ching-Long Lin (lung research, Univ. of Iowa), Gro Amdam (honey bee brain research, Arizona State Univ.), and researchers locally at Texas A&M such as Thierry Lints (songbird brain research).

**Broader application of our framework:** The rapid tracing and reconstruction algorithms and the validation framework developed in this project are expected to be applicable to other high-throughput microscopy methods. Chatter in the cutting process is common to all physical sectioning approaches involving a microtome, including Array Tomography, SBF-SEM, and ATLUM, thus the modeling framework we propose here will have a general utility beyond this project.

**Project timeline:** The project will be carried out according to the following timeline:

- Year 1, 1st half: §C.1 (complete task); §C.2 (prototype); §C.3 (prototype)
- Year 1, 2nd half: §C.1 (organize data into database); §C.2 (start testing on KESM data); §C.3 (compare results on unvalidated KESM reconstruction); Release initial raw data.
- Year 2, 1st half: §C.2 (large scale reconstruction); §C.3 (initial development of the SLICE module); §C.4 (initial model); §C.5 (selected manual validation); Release partial reconstruction.
- Year 2, 2nd half: §C.2 (revise reconstruction algorithm based on §C.5 from above); §C.3 (improved SLICE module based on §C.4 from above; start developing vascular development model); §C.4 (finalize KESM sectioning and imaging model); §C.5 (use the digital phantoms from the prototype SLICE module). Release partial reconstruction (enhanced and validated). Hold high school contest.
- Year 3, 1st half: §C.2 (continued enhancement of the reconstruction algorithm based on §C.5 from above); §C.3 (finalize SLICE module); §C.5 (generate full digital phantoms)
- Year 3, 2nd half: §C.5 (large scale validation); Release data, reconstruction, and reconstruction software.

## REFERENCES CITED

- [1] Abbott, L. C., and Sotelo, C. (2000). Ultrastructural analysis of catecholaminergic innervation in weaver and normal mouse cerebellar cortices. *Journal of Comparative Neurology*, 426:316–329.
- [2] Al-Kofahi, K. A., Lasek, S., Szarowski, D. H., Pace, C. J., Nagy, G., Turner, J. N., and Roysam, B. (2002). Rapid automated three-dimensional tracing of neurons from confocal image stacks. *IEEE Transactions on Information Technology in Biomedicine*, 6:171–187.
- [3] Allen Institute for Brain Science, Allen brain atlas. <http://www.brain-map.org/>.
- [4] Anderson, A. R. A., and Chaplain, M. A. J. (1998). Continuous and discrete mathematical models of tumor-induced angiogenesis. *Bulletin of Mathematical Biology*, 60:857–900.
- [5] Ascoli, G. (1999). Progress and perspectives in computational neuroanatomy. *The Anatomical Record (New Anatomy)*, 257:195–207.
- [6] Ascoli, G., Krichmar, J., Scorcioni, R., Nasuto, S., and Senft, S. (2001). Computer generation and quantitative morphometric analysis of virtual neurons. *Anatomical Embryology*, 204:283–301.
- [7] Ascoli, G. A. (2003). Variability of dendritic structure among and within morphological classes and experimental protocols. Presented during the Images as Phenotypes Satellite Symposium at the Society for Neuroscience 33rd Annual Meeting, New Orleans, LA.
- [8] Ascoli, G. A., and Krichmar, J. L. (2000). L-Neuron: A modeling tool for the efficient generation and parsimonious description of dendritic morphology. *Neurocomputing*, 32–33:1003–1011.
- [9] Boldak, C., Rolland, Y., and Toumoulin, C. (2003). An improved model-based vessel tracking algorithm with application to computed tomography angiography. *Journal of Biocybernetics and Biomedical Engineering*, 3:41–64.
- [10] Braitenberg, V., and Schuz, A. (1998). *Cortex: Statistics and Geometry of Neuronal Connectivity*. Berlin: Springer. Second edition.
- [11] Butz, M., and van Ooyen, A. (2006). Modelling structural plasticity and neural development. In *Connectionists: CNS\*2006: workshop on structural plasticity and development*.
- [12] Can, A., Shen, H., Turner, J. N., Tanenbaum, H. L., and Roysam, B. (1999). Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms. *IEEE Transactions on Information Technology in Biomedicine*, 3:125–138.
- [13] Cannon, R. C., Turner, D. A., Pyapali, G. K., and Wheal, H. V., Duke-Southampton archive of neuronal morphology. <http://neuron.duke.edu/cells/index/topindex.html>.
- [14] Cannon, R. C., Turner, D. A., Pyapali, G. K., and Wheal, H. V. (1998). An on-line archive of reconstructed hippocampal neurons. *Journal of Neuroscience Methods*, 84:49–54.
- [15] Capowski, J. J., editor (1989). *Computer Techniques in Neuroanatomy*. Plenum.
- [16] Cell Centered Database. <http://ncmir.ucsd.edu/CCDB/>.

- [17] Choe, Y., Abbott, L. C., Han, D., Huang, P.-S., Keyser, J., Kwon, J., Mayerich, D., Melek, Z., and McCormick, B. H. (2008). Knife-edge scanning microscopy: High-throughput imaging and analysis of massive volumes of biological microstructures. In Rao, A. R., and Cecchi, G., editors, *High-Throughput Image Reconstruction and Analysis: Intelligent Microscopy Applications*. Boston, MA: Artech House. In press.
- [18] Collins, D. L., Zijdenbos, A. P., Kollokian, V., Sled, J. G., Kabani, N. K., Holmes, C. J., and Evans, A. C. (1998). Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging*, 17:463–468.
- [19] Denk, W., and Horstmann, H. (2004). Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biology*, 19:e329.
- [20] Eberhard, J. P., Wanner, A., and Wittum, G. (2006). NeuGen: A tool for the generation of realistic morphology of cortical neurons and neural networks. *Neurocomputing*, 70:327–342.
- [21] Eshera, M. A., and Fu, K. S. (1986). An image understanding system using attributed symbolic representation and inexact graph-matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:604–618.
- [22] Fiala, J. C. (2005). Reconstruct: A free editor for serial section microscopy. *Journal of Microscopy*, 218:52–61.
- [23] Fiala, J. C., and Harris, K. M. (2001). Extending unbiased stereology of brain ultrastructure to three-dimensional volumes. *J. Am. Med. Inform. Assoc.*, 8:1–16.
- [24] Frangi, A. F., Niessen, W. J., Hoogeveen, R. M., Walsum, T. V., and Viergever, M. A. (1999). Model-based quantification of 3-D magnetic resonance angiographic images. *IEEE Transactions on Medical Imaging*, 18:946–956.
- [25] Gleeson, P., Steuber, V., and Silver, R. (2007). NeuroConstruct: A tool for modeling networks of neurons in 3D space. *Neuron*, 54:219–235.
- [26] Guntupalli, J. S. (2007). *Physical Sectioning in 3D Biological Microscopy*. Master's thesis, Department of Computer Science, Texas A&M University.
- [27] Han, D., Keyser, J., and Choe, Y. (2009). A local maximum intensity projection tracing of vasculature in Knife-Edge Scanning Microscope volume data. In *Proceedings of the IEEE International Symposium on Biomedical Imaging*. In press.
- [28] Hayworth, K. (2008). Automated creation and SEM imaging of Ultrathin Section Libraries: Tools for large volume neural circuit reconstruction. In *Society for Neuroscience Abstracts*. Washington, DC: Society for Neuroscience. Program No. 504.4.
- [29] Hayworth, K., and Lichtman, J. W. Automatic Tape-Collecting Lathe Ultramicrotome (ATLUM), [http://www.mcb.harvard.edu/lichtman/ATLUM/ATLUM\\_web.htm](http://www.mcb.harvard.edu/lichtman/ATLUM/ATLUM_web.htm).
- [30] Jacobowitz, D. M., and Abbott, L. C. (1997). *Chemoarchitectonic Atlas of the Developing Mouse Brain*. Boca Raton, FL: CRC Press.
- [31] Jain, V., Murray, J. F., Roth, F., Turaga, S., Zhigulin, V., Briggman, K. L., Helmstaedter, M. N., Denk, W., and Seung, H. S. (2007). Supervised learning of image restoration with convolutional networks. In *IEEE 11th International Conference on Computer Vision (ICCV 2007)*, 1–8.

- [32] Jurrus, E., Tasdizen, T., Koshevoy, P., Fletcher, P. T., Hardy, M., Chen, C.-B., Denk, W., and Whitaker, R. (2006). Axon tracking in serial block-face scanning electron microscopy. In Metaxas, D., Whitaker, R., Rittscher, J., and Sebastian, T., editors, *Proceedings of the 1st Workshop on Microscopic Image Analysis with Applications in Biology*, 114–119.
- [33] Kobayashi, A. (1996). *Machining of Plastics*. New York, NY: McGraw-Hill.
- [34] Koene, R. A. (2007). Large scale high resolution network generation: Producing known validation sets for serial reconstruction methods that use histological images of neural tissue. In *International Conference on Complex Systems*. [Presentation].
- [35] Koene, R. A., Tijms, B., van Hees, P., Postma, F., de Ridder, A., Ramakers, G. J. A., van Pelt, J., and van Ooyen, A. (2009). NETMORPH: A framework for the stochastic generation of large scale neuronal networks with realistic neuron morphologies. *Neuroinformatics*, 7:1539–2791.
- [36] Kwon, J., Mayerich, D., Choe, Y., and McCormick, B. H. (2008). Lateral sectioning for knife-edge scanning microscopy. In *Proceedings of the IEEE International Symposium on Biomedical Imaging*, 1371–1374.
- [37] Kwon, J.-R. (2009). *Acquisition and Mining of the Whole Mouse Brain Microstructure*. PhD thesis, Department of Computer Science, Texas A&M University.
- [38] MacKenzie-Graham, A., Jones, E. S., Shattuck, D. W., Dinov, I. D., Bota, M., and Toga, A. W. (2003). The informatics of a C57BL/6J mouse brain atlas. *Neuroinformatics*, 1:397–410.
- [39] Martone, M. E., Gupta, A., Wong, M., Qian, X., Sosinsky, G., Ludscher, B., and Ellisman, M. H. (2002). A cell-centered database for electron tomographic data. *J. Struct. Biol.*, 138:145–155.
- [40] Masood, S., Gao, J., and Yang, G.-Z. (2002). Virtual tagging: Numerical considerationsn and phantom validation. *IEEE Transactions on Medical Imaging*, 21:1123–1131.
- [41] Mayerich, D., Abbott, L. C., and Keyser, J. (2008). Visualization of cellular and microvessel relationship. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of IEEE Visualization)*, 14:1611–1618.
- [42] Mayerich, D., Abbott, L. C., and McCormick, B. H. (2008). Knife-edge scanning microscopy for imaging and reconstruction of three-dimensional anatomical structures of the mouse brain. *Journal of Microscopy*, 231:134–143.
- [43] Mayerich, D., and Keyser, J. (2008). Filament tracking and encoding for complex biological networks. In *Proceedings of ACM Symposium on Solid and Physical Modeling*, 353–358.
- [44] Mayerich, D., Kwon, J., Choe, Y., Abbott, L., and Keyser, J. (2008). Constructing high-resolution microvascular models. In *Proceedings of the 3rd International Workshop on Microscopic Image Analysis with Applications in Biology (MIAAB 2008)*. In press.
- [45] Mayerich, D., McCormick, B. H., and Keyser, J. (2007). Noise and artifact removal in knife-edge scanning microscopy. In *Proceedings of the IEEE International Symposium on Biomedical Imaging*, 556–559.
- [46] Mayerich, D. M., and Keyser, J. (2008). Filament tracking and encoding for complex biological networks. In *Proceedings of Solid Modeling*. To appear.

- [47] McCormick, B. H. (2003). The knife-edge scanning microscope. Technical report, Department of Computer Science, Texas A&M University. <http://research.cs.tamu.edu/bnl/>.
- [48] McCormick, B. H., System and method for imaging an object. USPTO patent #US 6,744,572 (for Knife-Edge Scanning; 13 claims).
- [49] McCormick, B. H., Abbott, L. C., Mayerich, D. M., , Keyser, J., Kwon, J., Melek, Z., and Choe, Y. (2006). Full-scale submicron neuroanatomy of the mouse brain. In *Society for Neuroscience Abstracts*. Washington, DC: Society for Neuroscience. Program No. 694.5. Online.
- [50] McCormick, B. H., Abbott, L. C., Mayerich, D. M., , Keyser, J., Kwon, J., Melek, Z., and Choe, Y. (2006). Full-scale submicron neuroanatomy of the mouse brain. In *Society for Neuroscience Abstracts*. Washington, DC: Society for Neuroscience. Program No. 694.5. Online.
- [51] McCormick, B. H., Mayerich, D., and Wiercigroch, M. (2005). Nanomachining for high-resolution scanning of mammalian brain microstructure. In *Proceedings of the 11th International Conference on Fracture*.
- [52] McCormick, B. H., and Mayerich, D. M. (2004). Three-dimensional imaging using Knife-Edge Scanning Microscope. *Microscopy and Microanalysis*, 10 (Suppl. 2):1466–1467.
- [53] McLachlan, G. J., and Krishnan, T. (1996). *The EM Algorithm and Extensions*. John Wiley and Sons.
- [54] Melek, Z., Mayerich, D., Yuksel, C., and Keyser, J. (2006). Visualization of fibrous and thread-like data. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1165–1172.
- [55] Melnik, S., Garcia-Molina, H., and Rahm, E. (2002). Similarity flooding: A versatile graph matching algorithm and its application to schema matching. In *Proceedings of the 18th International Conference on Data Engineering*, 117–128.
- [56] Micheva, K., and Smith, S. J. (2007). Array tomography: A new tool for imaging the molecular architecture and ultrastructure of neural circuits. *Neuron*, 55:25–36.
- [57] MicroBrightField, Inc., Autoneuron. <http://www.mbfbioscience.com/autoneuron>.
- [58] Miikkulainen, R., Bednar, J. A., Choe, Y., and Sirosh, J. (2005). *Computational Maps in the Visual Cortex*. Berlin: Springer. URL: <http://www.computationalmaps.org>.
- [59] Mikula, S., Trotts, I., Stone, J. M., and Jones, E. G. (2007). Internet-enabled high-resolution brain mapping and virtual microscopy. *Neuroimage*, 35:9–15.
- [60] Nakagami, H., Nakagawa, N., Takeya, Y., Kashiwagi, K., Ishida, C., ichiro Hayashi, S., Aoki, M., Matsumoto, K., Nakamura, T., Ogihara, T., and Morishita, R. (2006). Model of vasculogenesis from embryonic stem cells for vascular research and regenerative medicine. *Hypertension*, 48:112–119.
- [61] Pham, D. L., Xu, C., and Prince, J. L. (2000). Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2:315–337.
- [62] Pizlo, Z. (2001). Perception viewed as an inverse problem. *Vision Research*, 41:3145–3161.

- [63] Pluim, J. P. W., Maintz, J. B. A., and Viergever, M. A. (2000). Image registration by maximization of combined mutual information and gradient information. In *Lecture Notes In Computer Science, Vol. 1935; Proceedings of the Third International Conference on Medical Image Computing and Computer-Assisted Intervention*, 452–461. London: Springer.
- [64] Rhyu, I. J., and D. B. Walker, L. C. A., and Sotelo, C. (1999). An ultrastructural study of granule cell/Purkinje cell synapses in tottering ( $tg/tg$ ), leaner ( $tg^{la}/tg^{la}$ ) and compound heterozygous tottering/leaner ( $tg/tg^{la}$ ) mice. *Neuroscience*, 90:717–728.
- [65] Rosen, G. D., Williams, A. G., Capra, J. A., Connolly, M. T., Cruz, B., Lu, L., Airey, D. C., Kulakarni, A., and Williams, R. W. (2000). The mouse brain library@[www.mbl.org](http://www.mbl.org). In *Proceedings of the 14th International Mouse Genome Meeting, Narita, Japan*, C6 [online]. [http://www.imgs.org/abstracts/2000abstracts/C6\\_p039E\\_RobWilliams.shtml](http://www.imgs.org/abstracts/2000abstracts/C6_p039E_RobWilliams.shtml), <http://mbl.org/>.
- [66] Segev, R., and Ben-Jacob, E. (2000). Generic modeling of chemotactic based self-wiring of neural networks. *Neural Networks*, 13:185–199.
- [67] Sporns, O., Tononi, G., and Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Computational Biology*, 1:e42.
- [68] Tong, S., and Yuan, F. (2001). Numerical simulations of angiogenesis in the cornea. *Microvascular Research*, 61:14–27.
- [69] Tsai, P. S., Friedman, B., Ifarraguerri, A. I., Thompson, B. D., Lev-Ram, V., Schaffer, C. B., Xiong, Q., Tsien, R. Y., Squier, J. A., and Kleinfeld, D. (2003). All-optical histology using ultrashort laser pulses. *Neuron*, 39:27–41.
- [70] van Ooyen, A., van Pelt, J., and Corner, M. (1995). Implications of activity dependent neurite outgrowth for neuronal morphology and network development. *Journal of Theoretical Biology*, 172:63–82.
- [71] van Pelt, J., Schierwagen, A., and Uylings, H. (2001). Modeling dendritic morphological complexity of deep layer cat superior colliculus neurons. *Neurocomputing*, 38–40:403–408.
- [72] van Pelt, J., and Uylings, H. (1999). Modeling the natural variability in the shape of dendritic trees: Application to basal dendrites of small rat cortical layer 5 pyramidal neurons. *Neurocomputing*, 26–27:305–311.
- [73] van Pelt, J., and Uylings, H. (2002). Branching rates and growth functions in the outgrowth of dendritic branching patterns. *Network: Computations in Neural Systems*, 13:261–281.
- [74] van Pelt, J., and Uylings, H. (2003). Growth functions in dendritic outgrowth. *Brain and Mind*, 4:51–65.
- [75] van Pelt, J., and Uylings, H. (2005). Natural variability in the geometry of dendritic branching patterns. In Lindsay, K., Poznanski, R., Reeke, G., Rosenberg, J., and Sporns, O., editors, *Modeling in the Neurosciences*, chapter 4. London: CRC Press. Second edition.
- [76] Warfield, S. J., Zou, K. H., and Wells, W. M. (2002). Validation of image segmentation and expert quality with expectation-minimization algorithm. In *Lecture Notes In Computer Science, Vol. 2488; Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention*, 298–306.

- [77] Wiercigroch, M., and Budak, E. (2001). Sources of nonlinearities, chatter generation and suppression in metal cutting. *Philosophical Transactions of the Royal Society of London, A*, 359:663–693.
- [78] Yoo, T., Ackerman, N. J., and Vannier, M. (2000). Toward a common validation methodology for segmentation and registration algorithms. In *Lecture Notes In Computer Science, Vol. 1935; Proceedings of the Third International Conference on Medical Image Computing and Computer-Assisted Intervention*, 422–431. London: Springer.