Informatics

Spatial Neighborhood Analysis and Comparison for Nanoscale Brain Structures

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Visual Computing

eingereicht von

Jakob Troidl, BSc
Matrikelnummer 01525071

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Univ.Prof. Dipl.-Ing. Dr.techn. M. Eduard Gröller, TU Wien
Mitwirkung: Dipl.-Ing. (FH) Dr. techn. Johanna Beyer, Harvard University

Wien, 4. Juni 2021

J.Troidl

Jakob Troidl

M. Eduard Gröller

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

TU WIEN Informatics

Spatial Neighborhood Analysis and Comparison for Nanoscale Brain Structures

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieur

in

Visual Computing

by

Jakob Troidl, BSc
Registration Number 01525071

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.Prof. Dipl.-Ing. Dr.techn. M. Eduard Gröller, TU Wien
Assistance: Dipl.-Ing. (FH) Dr. techn. Johanna Beyer, Harvard University

Vienna, 4th June, 2021

J. Troidl
Jakob Troidl

M. Eduard Gröller

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Erklärung zur Verfassung der Arbeit

Jakob Troidl, BSc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 4. Juni 2021

J. Troidl

Jakob Troidl

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Danksagung

Zuallererst möchte ich meiner Betreuerin Johanna Beyer dafür danken, dass sie mich während des gesamten Prozesses unterstützt und mir so viel beigebracht hat. Danke, dass du deine Leidenschaft für Datenvisualisierung mit mir geteilt hast und mir geduldig geholfen hast, als ich nicht mehr weitergekommen bin und überfordert war! Das hat diese Zeit lustig und spannend gemacht.

Ich möchte auch Meister Eduard Gröller für seine Mentoring während meines ganzen Studiums danken. Du hast mich in die Welt der Forschung eingeführt. Ich weiß es zu schätzen, dass du mir geholfen hast, mein Studium individuell zu gestalten und dass du meine Pläne immer unterstützt hast.

Mein Dank gilt auch Markus Hadwiger, der ein großartiger Kollege, Mentor und Unterstützer war. Es ist großartig zu sehen, wie begeistert du und deine Gruppe von eurer Forschung seid. Das ist eine echte Inspiration für mich!

Ich möchte auch Hanspeter Pfister dafür danken, dass er mich in seine Forschungsgruppe an der Harvard University eingeladen hat, wo ich ein lebendiges intellektuelles Umfeld erlebt habe und viele inspirierende Forscher treffen durfte.

Ich habe das Glück, großartige Freunde zu haben, die mein Studium zu einer schönen Erfahrung gemacht haben und mich unterstützt haben, wenn ich Hilfe gebraucht habe. Danke, Valentin, Fabi, Max und Scholzi, dass wir zusammen die ersten Semester an der TU Wien bestritten haben. PK ohne euch wäre schrecklich gewesen. Ein großes Lob an Johannes und Basti, die mich ein Dutzend Mal aus meinem Zimmer geholt und zum Chillen gezwungen haben. Danke, Lara und Claas, meine Freunde in Boston, dass ihr die Zeit in den USA so prägend und schön gemacht habt, auch wenn COVID-19 nur ein paar Wochen nach unserer Ankunft losging.

Am meisten möchte ich mich bei meinen Eltern Robert und Stefanie, meinem Bruder Johannes, meinen Großeltern Burga und Heiner und meiner Partnerin Rubez bedanken. Danke, dass ihr Fragen gestellt habt, mich die Dinge aus einer anderen Perspektive sehen lasst und einfach da wart.

Dieses Projekt wurde von der Österreichischen Marshallplan-Stiftung, dem International Office der TU Wien und der King Abdullah University of Science & Technology (KAUST) finanziell unterstützt.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Acknowledgements

First and foremost, I want to thank my supervisor Johanna Beyer for supporting me throughout the process and teaching me so much. Thank you for sharing your passion for data visualization with me and for helping me patiently when I was stuck and overwhelmed! It made the whole process fun and exciting.

I also wish to thank Meister Eduard Gröller for his mentorship during my undergraduate and graduate studies and for introducing me to the world of research. I appreciate you helping me to shape my studies individually and for always supporting my plans.

I also want to thank Markus Hadwiger for being a great collaborator, mentor, and supporter. It is fantastic to see you and your group being so enthusiastic about research. This is a true inspiration to me!

Furthermore, I am also thankful to Hanspeter Pfister for inviting me to his research group at Harvard University, where I could experience a vibrant intellectual environment and meet many inspiring researchers.

I am fortunate to have great friends who made my studies a fun experience and lifted me up when I needed it. Thank you, Valentin, Fabi, Max, and Scholzi, for rocking through the first semesters as we started at TU Wien. PK without you would have been horrific. Big shout out to Johannes and Basti for getting me out of my study room a dozen times and forcing me to chill out. Thank you, Lara and Claas, my Boston buddies, for making my time in the US so formative and fun, even though COVID-19 hit just a couple of weeks after we arrived.

Most of all I want to thank my parents Robert and Stefanie, my brother Johannes, my grandparents Burga, Heiner and my partner Rubez. Thank you for asking questions, letting me see things from another perspective, and just being present.

This project was financially supported by the Austrian Marshall Plan Foundation, the International Office at TU Wien, and King Abdullah University of Science & Technology (KAUST).

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Kurzfassung

Neueste hochauflösende elektronenmikroskopische Aufnahmen ermöglichen Neurowissenschaftlern, nicht nur Zellen, sondern auch einzelne Zellstrukturen aus Mäusegehirnen zu rekonstruieren. Auf Basis dieser Daten erhoffen sie sich, ein besseres Verständnis der Gehirnfunktion und -entwicklung zu gewinnen, indem sie lokale Nachbarschaften von Zellstrukturen detailliert analysieren. Genaue Analysen erfordern jedoch einen effizienten und skalierbaren Vergleich unterschiedlicher Zellstrukturen, die von zwei bis zu Hunderten von lokalen Nachbarschaften reichen. Wissenschaftler müssen in der Lage sein, die 3D-Morphologie der Zellstrukturen, ihre räumlichen Verteilungen und Abstände sowie deren Korrelationen zu analysieren.

Die zentrale Prämisse dieser Arbeit ist, die Schwierigkeit eine einheitliche Lösung für die vergleichende Visualisierung bereitzustellen, um die gegebene breite Palette von Aufgaben zu unterstützen.

Deshalb haben wir NeuroKit als leicht konfigurierbaren Werkzeugkasten entwickelt, der es Wissenschaftlern ermöglicht, den Arbeitsablauf, die Visualisierungen und die unterstützten Benutzerinteraktionen an ihre spezifischen Aufgaben und Fragestellungen anzupassen. Darüber hinaus bietet NeuroKit einen skalierbaren, vergleichenden Visualisierungsansatz für die räumliche Nachbarschaftsanalyse von Gehirnstrukturen.

NeuroKit unterstützt interaktive 3D-Renderings sowie abstrakte quantitative Visualisierungen und ordnet diese automatisch in der Benutzeroberfläche an. Um neue domänenspezifische Analysen zu unterstützen, erlauben wir die Definition von individualisierten Visualisierungen und deren Parametern. Diese Konfigurierbarkeit ist mit einem neuartigen skalierbaren visuellen Vergleichsansatz verknüpft, der die Visualisierungen automatisch an die Anzahl der zu vergleichenden Strukturen anpasst.

Wir demonstrieren einen detaillierten Anwendungsfall für die Analyse von Mitochondrien und analysieren die Nützlichkeit von NeuroKit in einer qualitativen Studie mit Neurowissenschaftlern.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Abstract

Recent high-resolution electron microscopy imaging allows neuroscientists to reconstruct not just entire cells but individual cell substructures (i.e., cell organelles) as well. Based on these data, scientists hope to get a better understanding of brain function and development through detailed analysis of local organelle neighborhoods. However, in-depth analyses require efficient and scalable comparison of a varying number of cell organelles, ranging from two to hundreds of local spatial neighborhoods. Scientists need to be able to analyze the 3D morphologies of organelles, their spatial distributions and distances, and their spatial correlations. This thesis's central premise is that it is hard to provide a one-size-fits-all comparative visualization solution to support the given broad range of tasks and scales.

To address this challenge, we have designed NeuroKit as an easily configurable toolkit that allows scientists to customize the tool's workflow, visualizations, and supported user interactions to their specific tasks and domain questions. Furthermore, NeuroKit provides a scalable comparative visualization approach for spatial neighborhood analysis of nanoscale brain structures.

NeuroKit supports small multiples of spatial 3D renderings as well as abstract quantitative visualizations, and arranges them in linked and juxtaposed views. To adapt to new domain-specific analysis scenarios, we allow the definition of individualized visualizations and their parameters for each analysis session. This configurability is tied in with a novel scalable visual comparison approach that automatically adjusts visualizations based on the number of structures that are being compared. We demonstrate an in-depth use case for mitochondria analysis in neuronal tissue and analyze the usefulness of NeuroKit in a qualitative user study with neuroscientists.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Contents

Kurzfassung **xi**

Abstract **xiii**

Contents **xv**

1 Introduction **1**
1.1 Motivation & Problem Statement 1
1.2 Aim of the Work 3
1.3 Contributions 3
1.4 Structure of the Thesis 4

2 Related Work **7**
2.1 Visualization for Neuroscience 7
2.2 Comparative Visualization 13
2.3 Spatial Neighborhood Analysis 15
2.4 Visualization Specification and Generation Frameworks 16

3 Biological Background **19**
3.1 Terminology 19
3.2 Connectomics Workflow 20
3.3 Data 24

4 Goal & Task Analysis **27**
4.1 Design Study Methodology 27
4.2 Domain Goals 28
4.3 Analysis Tasks 29
4.4 Requirements 31

5 NeuroKit Toolkit Design **33**
5.1 Custom Analysis Scenarios 33
5.2 Automatic User Guidance 35

6 Visual Neighborhood Analysis **39**

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

6.1 Scalable Comparison Framework 39
6.2 Visual Elements 40
6.3 Interaction 44

7 Techniques **45**
7.1 NeuroKit Architecture 45
7.2 Data Preprocessing 45
7.3 Implementation 49

8 Use Case: Mitochondria Analysis **61**
8.1 Synaptic Neighborhood of Mitochondria 61
8.2 Mitochondria Placement 65
8.3 Mitochondric Spine Coverage 68

9 Evaluation **73**
9.1 User Study: Mitochondria Analysis 73
9.2 Qualitative Feedback 77
9.3 Findings 77

10 Conclusions and Future Work **83**
10.1 Summary 83
10.2 Discussion 83
10.3 Future Work 84
10.4 Conclusion 85

List of Figures **87**

Bibliography **91**

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

CHAPTER 1

Introduction

1.1 Motivation & Problem Statement

Neurobiologists are acquiring ever-increasing amounts of high-resolution image data of the mammalian brain, aiming to tackle questions regarding brain physiology, the onset of diseases, and the emergence of conciousness. Connectomics, in particular, approaches these questions by looking at the detailed connectivity between neurons at a nanoscale resolution to extract the brain's wiring diagram [LD11]. However, mammalian brains are staggeringly complex, with tens of millions of interconnected neurons, dozens of cell substructures (i.e., cell organelles) per neuron, and billions of synapses. Furthermore, scientists are not just interested in a purely structural wiring diagram of neurons. They also want to analyze and compare local spatial neighborhoods around structures of interest and cell organelles. For example, to evaluate how differences in spatial arrangement might influence brain development and neuronal connectivity, scientists need to compare different biological structures, analyze distances and spatial correlations, and look at how certain features are spatially distributed. However, the specific details of this analysis (e.g., goals, target structures, measurements) vary for each scientist. Therefore, scientists need tools that support their general workflow and scale to their data sizes, and, at the same time, are flexible enough to allow for custom modifications to suit a scientist's individual workflow.

Most visualization tools for connectomics provide either 3D views for exploring large, segmented electron microscopy (EM) volumes [BAAK+13, BHAA+13, HBJP12], or focus on the analysis of synaptic connectivity [AABS+14, SBS+13]. While there is some work on analyzing the spatial distribution of glycogen granules in astrocytes [MAAB+18], little work has focused on the scalable visual comparative analysis of neuronal structures in high-resolution EM data. Furthermore, no integrated system can adapt to scientists' different, specific needs and domain goals. In fact, the particular workflow and analysis goals of neuroscientists often differ even between members of the same research group.

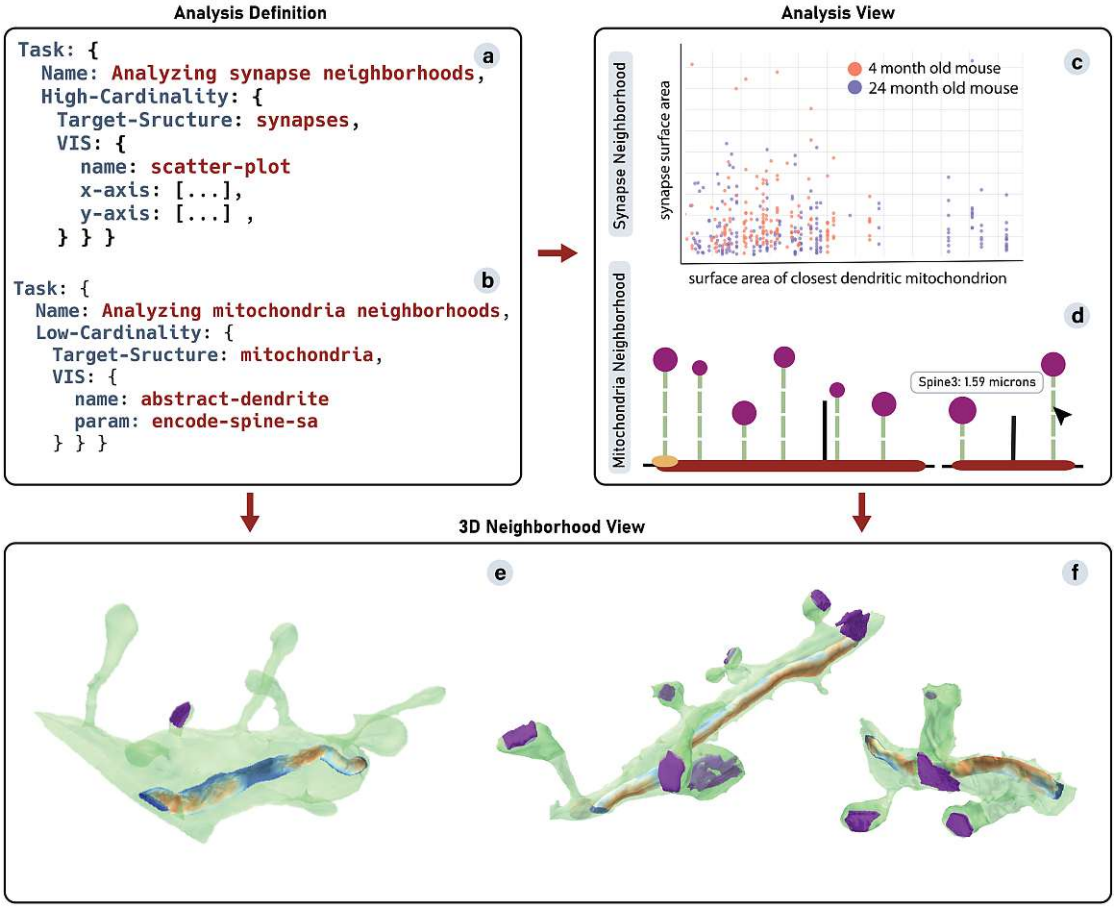


Figure 1.1: **Customized scalable comparison of spatial neighborhoods.** NeuroKit supports user-defined analysis scenarios (a,b) and automatically updates and adjusts visualizations and view parameters, to guide users in their specific analyses. Quantitative analysis views (c, d) enable scalable comparisons at different cardinality levels. Detailed 3D neighborhood views (e, f) depict the surroundings of structures of interest. Synapse domain task: (a) Synapse analysis scenario. (c) Scatterplot of hundreds of synapses from two different specimen. (e) Selected synapse (purple) and its neighborhood. Mitochondria domain task: (b) Mitochondria analysis scenario. (d) Distance-based abstraction of dendritic spines and their relation to mitochondria. (f) Mitochondria color-coded with their distance to the neurite membrane.

However, visualization tools designed to be general enough to support several types of analyses often result in overwhelming user interfaces and unintuitive workflows. On the other hand, tools designed for a specific analysis task typically do not generalize well to other tasks.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

1.2 Aim of the Work

In this work, we present NeuroKit, a novel tool for the interactive analysis of spatial neighborhoods in high-resolution 3D electron microscopy (3DEM) data, with the goal of providing scientists with a highly configurable visualization toolkit suitable to their domain goals.

First, NeuroKit is a general framework that is expressive and powerful, but can be customized to support individualized workflows. By allowing scientists to customize their analysis workflow to best suit their needs, we minimize visual complexity, streamline the user interface, and simplify user interaction (see Fig. 1.1). We have designed the NeuroKit framework in such a way that it is easily extensible. Scientists can specify new, individualized analysis scenarios and appropriate visualizations that support their analysis tasks. Second, we support the domain goals of our collaborating neuroscientists for analyzing and comparing spatial neighborhoods. We have designed a novel scalable comparison approach that automatically adjusts visualizations based on the number of structures that are being compared, and we support spatial and analytic reasoning by combining spatial 3D views with abstract quantitative views. Fig. 1.2 depicts the NeuroKit system design.

1.3 Contributions

The main contributions of this work are the following:

- First, a *customizable framework for spatial neighborhood analysis* of 3DEM data, facilitating the analysis of spatial distributions, distances, and 3D morphology. Our framework supports the typical 3DEM analysis workflow, but also allows scientists to specify custom analysis tasks and visualizations.

- Second, a *scalable visual comparison framework* for spatial structures that allows the comparison of a wide range of element cardinalities, e.g., ranging from two to hundreds of structures within a single comparison.

- Third, a *custom analysis scenario* implemented in NeuroKit to study mitochondria in neuronal tissue. We have identified this scenario with our neuroscience collaborators, and describe the design of novel visual encodings and their use in NeuroKit to analyze mitochondria, their locations within a cell, and their relation to synapses.

We validate and demonstrate the usefulness of NeuroKit in a user study with domain experts and a detailed expert use case.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

1.4 Structure of the Thesis

In Chapter 2, we compare this work to previous and related work, including visualization for connectomics, comparative visualization, and visualization specification and generation frameworks. This helps to categorize our project and puts it into a bigger context. Chapter 3 introduces the reader to the biological background. We give a brief introduction to the relevant terminology and familiarize the reader with each step of the connectomics workflow.

Chapter 4 conducts a detailed goal & task analysis of NeuroKit. We first explain the design study methodology used during this project and then present the findings in the form of domain goals, tasks, and requirements. In Chapter 5, we illustrate how we translated the results of Chapter 4 into a usable toolkit. Chapter 6 explains our scalable visual neighborhood analysis approach and how we facilitate visual comparison between a varying number of spatial neighborhood instances. We also explain interaction techniques supporting these goals.

In Chapter 7 we present the techniques used for implementation, such as an order-independent transparency algorithm, all necessary data preprocessing steps, and the design space of the analysis definition files used for the customization of NeuroKit.

Chapter 8 describes a detailed usage scenario of NeuroKit investigating spatial neighborhoods of mitochondria. We showcase NeuroKit by presenting custom visualizations and discuss the expert interviews held while designing this scenario. Additionally, we demonstrate its usefulness and effectiveness in a qualitative user study in Chapter 9.

Chapter 10 concludes this thesis by discussing and summarizing our work and presenting an outline of future work.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

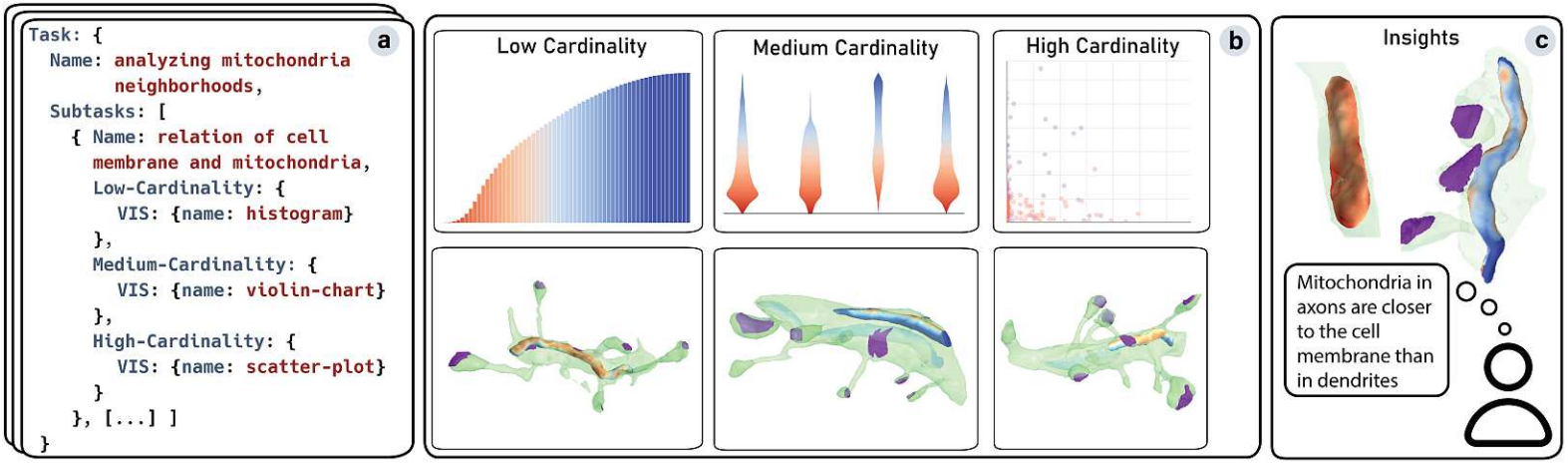


Figure 1.2: **NeuroKit design.** (a) Users can specify custom analysis scenarios in *analysis definition files*. Each user-defined *scenario* can consist of multiple *domain tasks* and incorporate customized visualizations and view parameters. (b) Our comparative visualization framework supports *low-cardinality*, *medium-cardinality*, or *high-cardinality* comparative analysis views (b, top row). Based on the number of structures in a comparison, NeuroKit automatically shows the corresponding view. To analyze complex 3D neighborhoods in more detail, we link quantitative views with 3D neighborhood views (b, bottom row). (c) Scientists can interactively refine the visual analysis to gain insight into their complex data sets.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 2

Related Work

In the following sections, we present related work to put this thesis into a bigger context. First, we offer recent work in visualization for neuroscience. Second, we analyze existing approaches in comparative visualization and delimit them from our work. Categorizing our approaches within spatial neighborhood analysis and visualization specification and generation frameworks is also essential to view our works' context.

2.1 Visualization for Neuroscience

Pfister et al. [PKB+12] categorize methods and challenges for the visualization of brain connectivity in macro-, meso-, and micro-scale analysis approaches (see Fig. 2.1). Here,

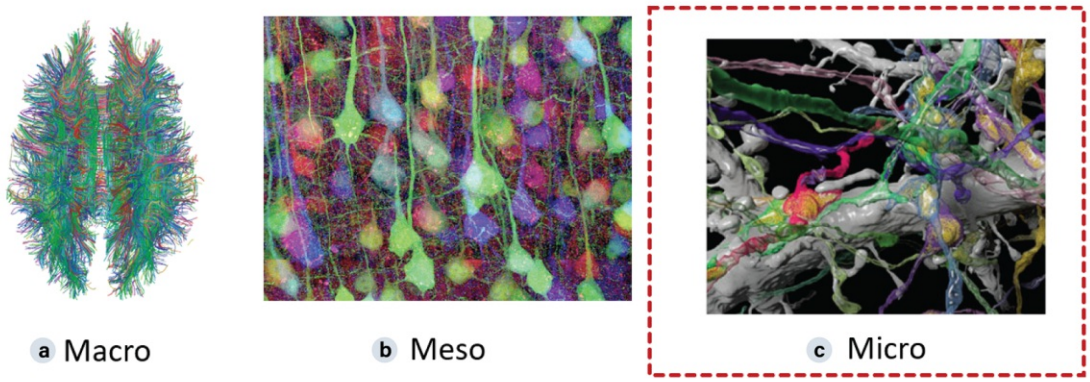
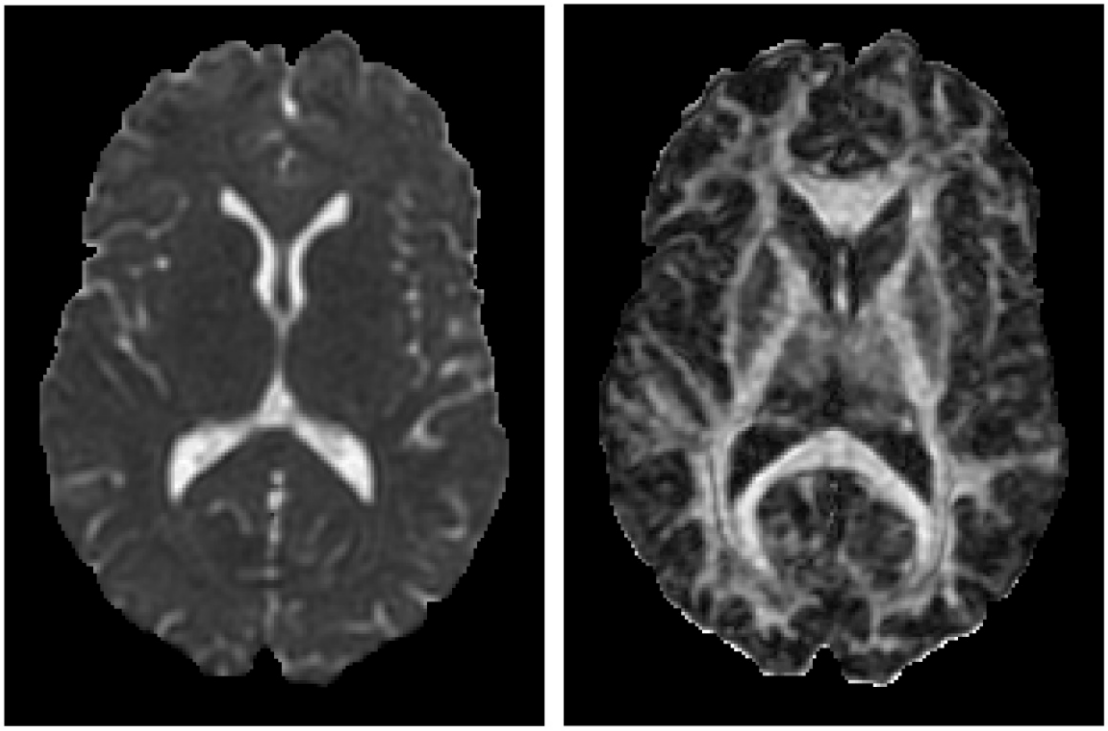


Figure 2.1: **Different scales of visualizing brain connectivity.** NeuroKit focuses on a micro-scale analysis (c). Source: (a) Macro [GHK+08], (b) Meso [Seu12], (c) Micro [PKB+12]

we describe related work of each of the three approaches in detail.

2.1.1 Macroscale Analysis

Macroscale analysis maps connectivity between larger brain regions such as the different brain lobes. This can help researchers to understand how brain regions interact and influence each other. This section will focus on visualization methods based on diffusion tensor imaging. DTI derives a second order diffusion tensor for each voxel describing water diffusion directions at this position in the brain. Vilanova et al. [VZKL06] classify DTI visualization approaches based on how the diffusion tensor field is reduced and whether the visualization shows local or global information. Thus, they derive five visualization categories. In this section, we describe each of these categories in detail. The following descriptions of the categorizations are based on Preim and Botha [PB13] and Vilanova et al. [VZKL06].



(a) Mean diffusivity (b) Fractional anisotropy

Figure 2.2: **Scalar visualization of DTI data** using mean diffusivity and fractional anisotropy. Source: [VZKL06].

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Scalar Reduction

Here, each 6D diffusion tensor is reduced to a scalar value (1D). Several methods have been proposed to map the 6D tensor to meaningful scalars. An example is the *mean diffusivity* $\langle D \rangle$ (see Fig. 2.2a)

$$\langle D \rangle = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}$$

whereby λ_i are the eigenvalues of the eigendecomposition of D. $\langle D \rangle$ quantifies the overall diffusion rate by computing the average of the eigenvalues of the diffusion tensor D. Another scalar index is the *fractional anisotropy* $\langle FA \rangle$ (see Fig. 2.2b)

$$\langle FA \rangle = \sqrt{\frac{3}{2}} \frac{\|D - \langle D \rangle I\|}{\|D\|}$$

with I as the identity matrix and $\|D\| = \sqrt{D : D}$ being the tensor norm of the mean diffusivity. FA also uses the eigenvalues of D. However, FA is relatively prone to noise. Since both measures are scalar values, they do not take the direction of the diffusion into account. Both, *mean diffusivity* and *fractional anisotropy* are invariant against rotations of the data set.

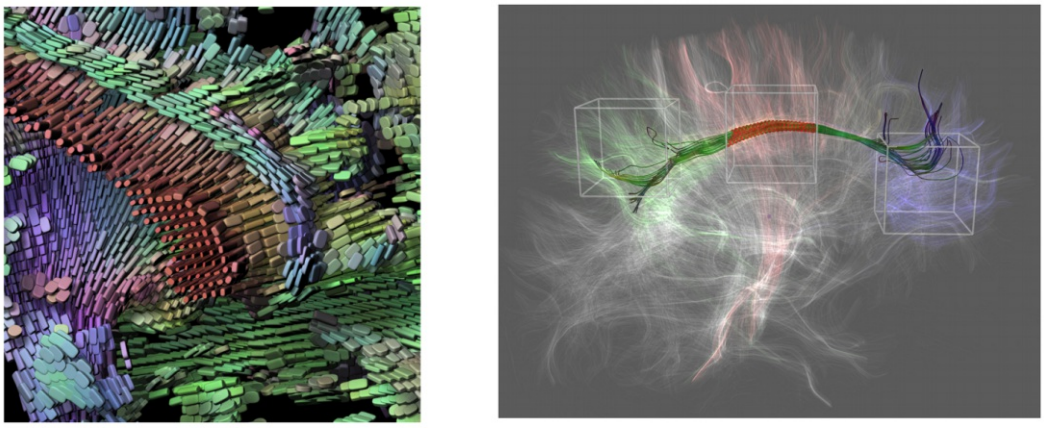
Volume Rendering

The previously presented visualizations of scalar fields were slice-based. Slice-based illustrations visualize 3D data by scrolling through the third dimension. Another method to visualize 3D scalar fields is *direct volume rendering*. By defining a transfer function assigning color and opacity to a scalar value, and tracing viewing rays through the volume, a 2D image is generated from a 3D data set. Kindlmann et al. [KWH00] propose specific transfer functions and shading techniques for volume renderings of DTI data sets. For instance, anisotropy shading is combined with standard Phong shading to illuminate the rendering.

Tensor Glyphs

Using tensor glyphs, no information of the diffusion tensor is discarded. Tensor glyphs are geometric objects with six degrees of freedom, such that they can fully represent a diffusion tensor. For instance, ellipsoids, superquadrics, or boxes are used as tensor glyphs. Figure 2.3a shows an illustration how diffusion tensors are visualized with superquadrics. Note that a glyph-based visualization does not encode correlations between diffusion tensors. Instead, these correlations are perceived by looking at multiple spatially close glyphs. Different methods have been proposed to arrange glyphs effectively in 3D space to optimize the perception of diffusion data. Laidlaw et al. [LAK+98] normalize the tensor glyphs to pack them more densely into 3D space. Additionally, they present a

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) Superquadrics (b) Streamlines

Figure 2.3: **DTI based connectivity visualizations.** Superquadrics visualizing the diffusion tensor per voxel. By tracing streamlines through the field of principal eigenvectors a full-brain tractography is derived. Source: [Kin04, BBP+05].

technique to color the glyphs by using brush strokes inspired by oil painters such that connectivity patterns become visible. In addition, they arrange glyphs in layers to create better depth perception. The biggest limitation of using tensor glyphs is that they are not clearly distinguishable in image space. For instance, the length of a superquadric glyph is not visible once it is oriented along the viewing direction.

Vector Field Visualization

By reducing each diffusion tensor to its principal eigenvector, standard vector field visualization techniques become applicable to the visualization of DTI data. Fiber-tracking [MvZ02] or tractography [BPP+00] traces streamlines through this vector field. Streamlines are computed by solving the following equation

$$p(t) = \int_0^t v(p(s))ds$$

with $p(t)$ being the streamline whereby $p(0)$ is its starting point. v referes to the vector field. In practice, this equation is numerically approximated and solved by using either Euler [Eul45] or Runge Kutta [Run95] integration. The streamlines are color-coded by mapping the direction of the vectors to a color. Figure 2.1 and Figure 2.3b show examples of streamlines traced through a field of principal eigenvectors.

Beyond Vector Field Visualization

By using more information than the principal eigenvector, researchers aim to reduce information loss when visualizing tensor fields. For instance, Zhang et al. [SDL03] use

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

stream surfaces to represent planar diffusion areas that can not be represented precisely with tubes or lines. Another approach by Hesselink et al. [HLL97] extracts topological skeletons from the field of diffusion tensors.

2.1.2 Mesoscale Analysis

Mesoscale analysis leverages light microscopy to image nerve cells and their substructures such as axons or dendrites. However, the resolution of light microscopy is not high enough to visualize synapses between nerve cells. Three main imaging techniques use light microscopy to visualize neuronal curcuits [PKB+12]. First, *single-cell staining by dye impregnation* uses the so-called Golgi stain to mark specific nerve cells within a dense block of nerve tissue. Second, scientists use *diffusion staining* [GGW+00] to trace nerve trajectories through brain regions. They inject a stain *in vivo* into a nerve cell which then diffuses along the path of the cell and therefore colors it. Third, the *brainbow* technique [LLS08] produces densely labeled images of brain tissue (see Fig. 2.1b) by breeding genetically modified mice, such that their brain tissue produces photophysical proteins. These proteins emit colored light under the microscope, and therefore nerve cells can be identified in the brain tissue.

2.1.3 Microscale Analysis

As the wavelength of light limits conventional microscopy, microscale analysis uses electron microscopy (EM) to image brain tissue at an even higher resolution. EM allows neuroscientists to visualize neuronal structures at a nanometer resolution and is, therefore, the only imaging method allowing the identification of individual synapses. Imaging a cubic millimeter of brain tissue using serial EM would produce data sets up to 800 TB of size [PKB+12]. These enormous amounts of data result in new computational challenges to extract connectivity information. Therefore, scientists usually work with much smaller volumes of brain tissue.

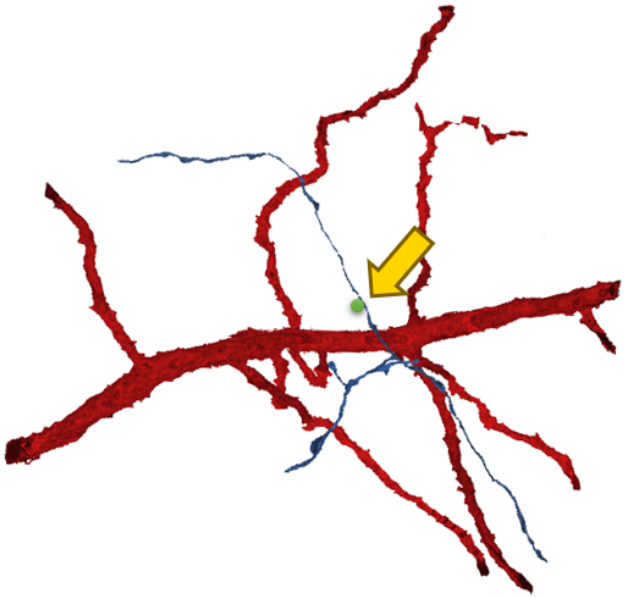
After densely labeling the EM image stack, 3D models of neuronal structures are reconstructed (see Fig. 2.1c, Sec. 3.2.3). These models are then used as the basis for a further detailed analysis.

This work focuses on micro-scale tissue visualization since we use EM data sets to resolve individual synapses to reconstruct 3D models of nerve cells. In the following, we use the term connectomics synonym to micro-scale brain tissue analysis even if our work does not specifically focus on brain connectivity.

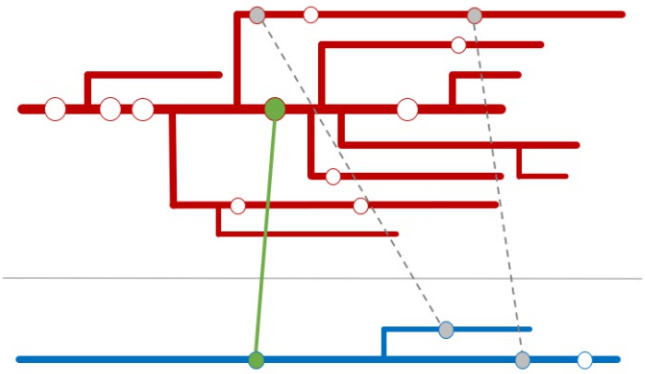
Visualization approaches have been applied to the entire connectomics data aquisition and analysis pipeline [HHM+17], ranging from visual segmentation tools [BSL18, GG90] and visual tools for proofreading of segmentations [AABH+16, HKT+18] to data exploration and analysis approaches [BAAK+13, AABS+14, MAAB+18, CBB+16].

Data exploration typically focuses either on the original *large-scale microscopy data* and its segmentation, or on *higher-level connectivity information* [PKB+12].

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) Two neurons rendered in 3D



(b) Neuron abstraction using a subway metaphor

Figure 2.4: **Micro-scale brain connectivity visualization** with Neurolines [AABS+14]. 3D rendering of two neurons forming a synapse. Skeletons of these two neurons abstracted using a subway metaphor. Links between the skeletons illustrate synapses.

Methods for structural connectivity analysis of neuronal cells often use abstract visual metaphors such as subway maps [AABS+14] or circuit diagrams [SBS+13] that highlight synaptic connections. Neurolines [AABS+14] allows viewing the connectivity between

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

neurons in an abstracted way. They reduce the three-dimensional nerve cells' complexity by abstracting them to two-dimensional straightened skeletons (see Fig. 2.4). They support visualizing skeletons at different scales to allow a simultaneous analysis of thousand of nerve cells. Sorger et al. [SBS+13] propose neuroMap to visualize nerve cell connectivities in the fruit fly brain as an interactive two-dimensional graph. In contrast, our work does not aim to highlight brain connectivity, but instead focuses on providing an insight into the spatial neighborhoods of neuronal structures and cell organelles.

On the other hand, methods for exploring the general morphology of the data typically focus on scalable visualization of the original or segmented electron microscopy data [HBJP12, BHAA+13]. For instance, ConnectomeExplorer [BAAK+13] supports the exploration of high-resolution EM data by integrating a domain-specific query language into a scalable volume visualization framework. Neurolucida [GG90], on the other hand, focuses on the segmentation of high-resolution EM data. Additionally, it allows neurobiologists to perform a quantitative analysis of single segmented structures, such as extracting neuron features like length, volume, or surface area. Similarly, Berger et al. [BSL18] present an interactive segmentation and annotation tool for high-resolution EM data, which can also be used for data inspection in 2D and 3D. While NeuroKit allows overlaying a 3D reconstruction with the original EM data, it is based on surface rendering techniques rather than direct volume visualization.

Other approaches focus more on the analysis of structures related to the brain's energy metabolism. Agus et al. [ACA+19] enable domain experts to investigate energy consumption in relation to brain morphology with a specific focus on glycogen granules. Mohammed et al. [MAAB+18] propose a visual tool to study interactions between neurites and glial cells. Thereby, they introduce a 2D visual abstraction panel to allow continuous user-specified transitions between abstraction levels. However, none of these approaches analyzes neighborhoods of selected neuronal structures and also does not support comparative visualizations of them.

2.2 Comparative Visualization

Comparative visualization approaches allow the direct comparison and assessment of similarities and differences between multiple data points or data sets. Pagendarm and Post [PP95] classify comparative visualization into *image-level* and *data-level* comparisons. *Image-level* comparison processes the data sources with different visualization pipelines and then compares the resulting images, while *data-level* comparison uses the same visualization pipeline to prepare the data for visual comparison. Our approach classifies as a data-level comparison method.

Gleicher et al. [GAW+11] identify three approaches to compare objects visually: juxtaposition, superposition, and explicitly encoding differences or similarities. Additionally, Kim et al. [KCK17] add interchangeability as a comparison method. Figure 2.5 shows an illustration of four different comparative visualization techniques. Juxtaposition displays entities side by side, while superimposition overlays them to facilitate their

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

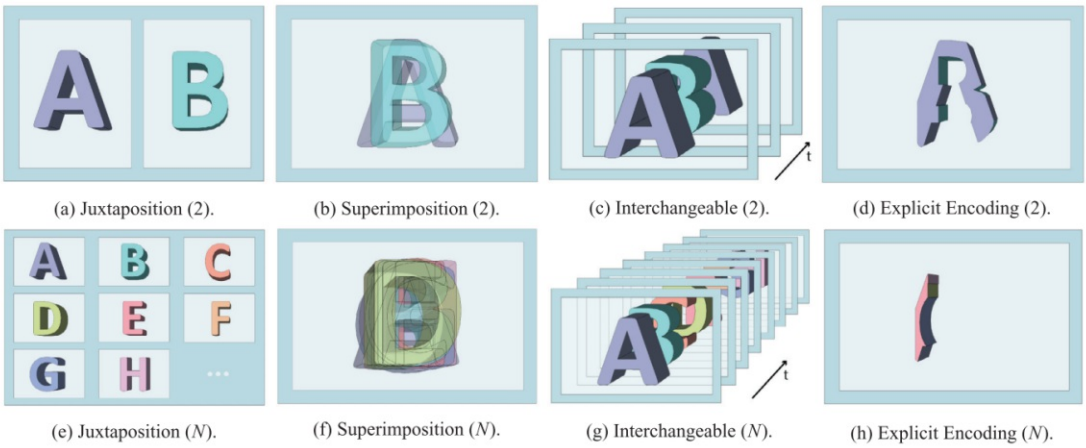


Figure 2.5: **Four different visual comparison approaches.** Each column shows a specific visual comparison approach (from left to right: Juxtaposition, Superimposition, Interchangeable, and Explicit Encoding). The top row indicates the comparison of two entities while the bottom row shows how the respective approach applies to comparing N entities. Source: [KCK17]

comparison. Both methods work well when comparing a low number of entities. However, both can result in cluttered visualizations as soon as the number of entities increases. Interchangeability relies on animations to visualize similarities or differences in objects. This method scales well, but a time-dependent visualization is often not suitable (e.g., can not be used in a book). Finally, explicit encoding processes multiple entities to display one composite of the different data sets. For instance, computing the intersection or difference between two 3D shapes is an example of explicit encoding.

Our approach uses a combination of juxtaposition and superimposition to compare neuronal structures and spatial neighborhoods. Comparing many structures simultaneously requires scalable visualization methods, such as small multiples for large-scale genome interaction matrices [LBK+18] or guaranteed visibility for tree comparisons [MGT+03]. Schmidt et al. [SPA+14] assess surface meshes by comparing each mesh against a reference mesh. For our data, reference meshes do not exist since the overall shapes and branching patterns of neurites vary significantly.

Most mesh comparison methods focus on meshes that are either (1) processed with different algorithms, (2) the output of multiple runs of the same algorithm with different parameters, or (3) results of numerical simulations with different parameter configurations. According to Kehrer et al. [KH13], category (1) falls into multi-model data, and (2) and (3) lie in the field of multi-run data. Our work lies in neither of those categories because the meshes of neuronal structures differ naturally and are not influenced by variations of computational models.

We identified five main classifications for the comparison of meshes. First, the number

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

of meshes that can be compared is different. In general, approaches can be subdivided into pairwise mesh comparison and multi mesh comparison methods. Second, the type of data source differs. Some approaches use measured data like meshes derived from point clouds or CT volume data. Another data source is simulation data/ensemble data. Third, some methods use information visualization techniques like parallel coordinates plots (PCP), scatter plots, or box plots, while other methods exclusively focus on the use of three-dimensional renderings. Fourth, most methods try to quantify the meshes using shape descriptors or similar methods. However, some methods only rely on visual comparison without computing deviation values or other numeric values. The fifth category concerns the ability of the visualization method to handle time-varying data.

For comparing scalar and flow fields, contour trees and graph matching have been used [HSKK01, SWC+08]. Hilaga et al. [HSKK01] describe a method for topology matching of 3D shapes using multiresolution Reeb graphs and use them as a search key to identify similarities in a database of 3D shapes. NeuroKit incorporates a tree-based comparison view, but focuses more on distances between structures in the tree rather than on the raw topology. Graph comparison techniques have also been proposed for analyzing brain connectivity [ABHR+13]. These techniques compare weighted graphs using node-link diagrams and augmented adjacency matrices. They further show the effectiveness of both visualizations in an evaluation with domain experts and conclude that matrix-based visualizations are better suited for brain connectivity analysis. Furthermore, Maries et al. [MMH+13] use visual comparison for spatial and non-spatial features of geriatric data sets. By combining medical imaging, mathematics, and interactive visualization methods, they can visualize and compare high-dimensional and multimodal data. Also, they use sparse partial least squares and Tikhonov regularization algorithms to quantify potential neurological diseases.

2.3 Spatial Neighborhood Analysis

Identifying and searching for patterns in biological tissue and cell neighborhoods is an active area of research. For highly multiplex imaging data, tools such as CytoMap [SFG+20] and Facetto [KBJ+20] have been proposed, which support the search for visual patterns in cell neighborhoods and hierarchial clustering of cell types for phenotype analysis, respectively.

Pettit et al. [Pet15] propose a method based on Hidden Markov random fields to cluster cells depending on their gene expression data. Although their approach was also evaluated on brain data, our primary goal is not to cluster cells but to visually analyze their surrounding area. Nhu et al. [NDBB17] present a novel tool to analyze several cells' interactions in three-dimensional data sets. Their solution is also capable of classifying cell types. While analyzing interactions between cells is closely related to our work, their method works on different data. They test it on Langerhans cells, which are dendritic cells in the epidermis and, therefore, not part of the central nervous system. Fernandez-Gonzalez et al. [FGBHOdS05] introduce a function to assess whether there

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

are patterns in a cell population. They solely focus on cellular relationships, while we also investigate the relationships between cells and cell organelles.

All of these approaches use multiplex imaging or microscopy data at a much lower resolution than EM data. Subsequently, the above approaches focus on larger cell neighborhoods (e.g., entire tissues or brain regions). NeuroKit, on the other hand, focuses on localized spatial neighborhoods around cell organelles. Jorstad et al. [JNC+15] use Blender and Neuromorph to allow users to manually extract quantitative measures from reconstructed meshes of cell organelles in high-resolution EM data. However, they focus on manual user interaction and simple measurements rather than neighborhood analysis and visual comparisons between organelles.

2.4 Visualization Specification and Generation Frameworks

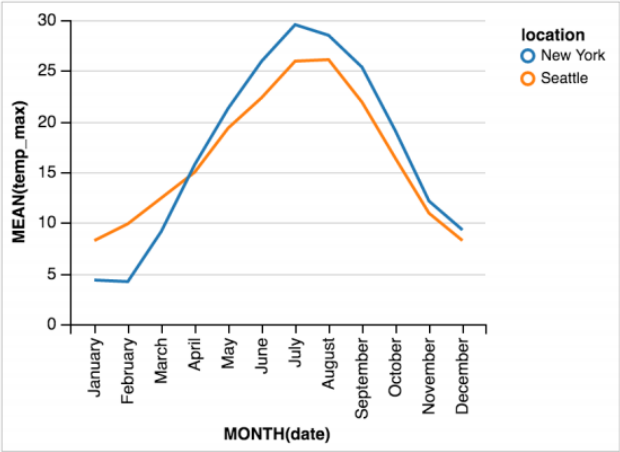
Frameworks for the specification and generation of visualizations can be classified into low level [BOH11, BH09] and high level [SWH14, SH14, SRHH16, SMWH17] approaches, based on the level of abstraction they provide in their interface. Many techniques use declarative methods (i.e., defining *what* not *how*) to specify data visualizations. Low-level visualization specifications like D3 [BOH11] require users to compose a visualization programmatically by specifying its components, while high-level approaches may support interactive data binding and visualization specification using drag and drop [SH14].

Vega-Lite [SMWH17], on the other hand, uses a grammar based on the JSON file format to specify charts. Figure 2.6 shows two visualization examples generated using Vega-Lite and their related JSON specifications. Methods like Lyra [SH14] allow complete interactive data binding and visualization specification using drag and drop. Furthermore, Satayanarayan et al. [SWH14] propose purely declarative methods to specify interactions with data visualizations. NeuroKit is inspired by these approaches, but offers an interface for *domain-specific visualization specification*. We allow users to define domain-specific analysis scenarios and domain tasks and map different visualizations to them. Our system parses this input and automatically creates customized visualizations.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

(a) Line chart with aggregation

```
{
  "data": {
    "url": "data/weather.csv",
    "formatType": "csv" },
  "mark": "line",
  "encoding": {
    "x": {
      "field": "date",
      "type": "temporal",
      "timeUnit": "month" },
    "y": {
      "field": "temp_max",
      "type": "quantitative",
      "aggregate": "mean" },
    "color": {
      "field": "location",
      "type": "nominal" }
  }
}
```



(b) Correlation between wind and temperature

```
{
  "data": {
    "url": "data/weather.csv",
    "formatType": "csv" },
  "mark": "point",
  "encoding": {
    "x": {
      "field": "temp_max",
      "type": "quantitative",
      "bin": true },
    "y": {
      "field": "wind",
      "type": "quantitative",
      "bin": true },
    "size": {
      "field": "*",
      "aggregate": "count" },
    "color": {
      "field": "location",
      "type": "nominal" }
  } }
```

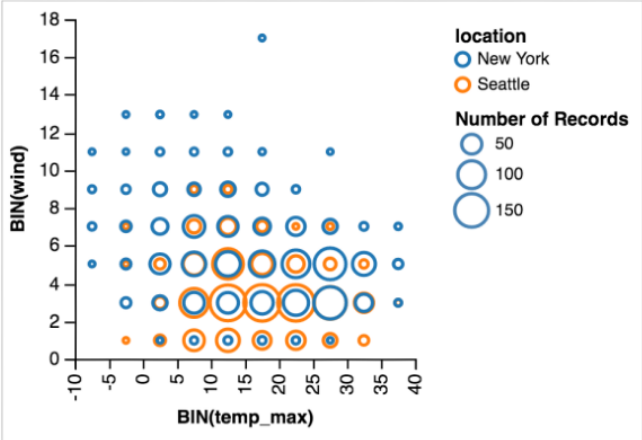


Figure 2.6: **Examples of visualizations generated with Vega-Lite.** (a) JSON specification of a line chart and the resulting visualization. (b) JSON specifiction of a scatter plot and the resulting visualization. Source: [SMWH17]

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

CHAPTER 3

Biological Background

Now, we explain the relevant concepts of neuroscience by giving a short introduction into the terminology, followed by an overview of the connectomics workflow. In the last section of this chapter, we present the data sets used for this work.

3.1 Terminology

The human brain consists of billions of interconnected nerve cells, or *neurons*. Neurons receive input on tree-like structures called *dendrites* and transmit signals via long tubular structures called *axons* (see Fig. 3.1). Axons and dendrites (collectively called *neurites*) connect via *synapses* to neighboring cells [Stu08]. The human brain consists of approximately tens of billions of neurons [ACG+09] that are interconnected through one hundred trillion synapses. The scientific field of connectomics aims to reconstruct this neural wiring diagram of the brain [8] to better understand the relation between the structure of the nervous system and its function [9]. Considering these vast number of structures and connections in the human brain, connectomics faces interesting challenges on both the biological and the computational side. In addition to synaptic connectivity, subcellular structures of neurons, so-called *cell organelles*, are also of high interest to neuroscientists. *Mitochondria*, for example, are cell organelles responsible for providing energy to the cell and its synapses [SW06]. Mitochondria can vary significantly in shape and are found in the cell body and axons, and dendrites. Other cell organelles, for example, are ribosomes or endoplasmic reticula. While ribosomes synthesize proteins, endoplasmatic reticula are responsible for various tasks such as hormone synthesis or storage of carbohydrates and calcium. Since cell organelles play an important role for metabolism and cell development, neuroscientists want to understand how they are distributed within neuronal structures and how they influence their spatial neighborhoods.

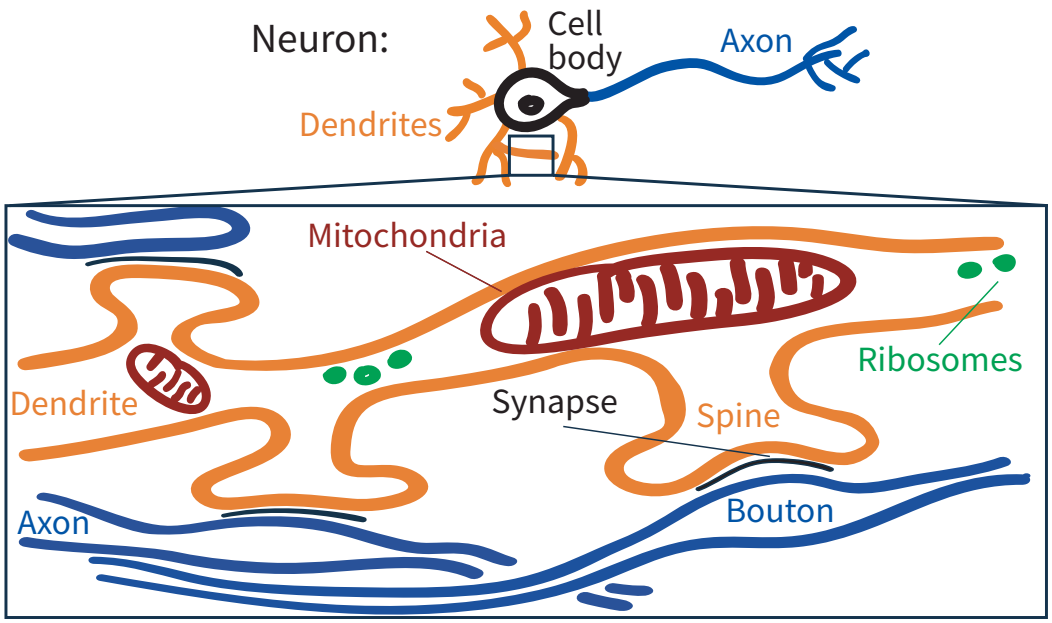


Figure 3.1: **Structure of neurons and cell organelles.** Nerve signals are transmitted from axons to dendrites via synapses. Neurons contain several different cell organelles, such as mitochondria or ribosomes. Mitochondria are responsible for providing energy to the neuron while ribosomes synthesize proteins. Source: Johanna Beyer

3.2 Connectomics Workflow

Connectomics aims to reconstruct the neural wiring diagram of the brain to understand how the its morphology relates to its function [Seu12]. Before analyzing the connectomics workflow, we highlight some challenges that must be addressed when working with high-resolution data of brain tissue. Lichtman and Denk [LD11] summarize these challenges that researchers currently face. First, the brain tissue consists of a multitude of different cell types. Neuronal structures are highly complex in their morphology and 3D shape. This makes it hard to classify them and to map their shape to its function. Second, brain function is not solely determined by the wiring diagram of its nerve cells but also by chemical and electrical activity. Observing chemical and electrical activity for many different cells is difficult, which adds another level of complexity. Third, branches of nerve cells span large volumes. For instance, dendrites of pyramidal cells extend through volumes of one cubic millimeter. Since connectomics requires data resolutions of several nanometers, a single and complete cell requires enormous amounts of data storage. Traditional light microscopes are not suited to achieve a such high resolution. Therefore, scientists use electron microscopes (EM) to image brain tissue at a nanometer resolution. However, EM imaging requires additional and time-consuming probe preparations. Finally, all structures in EM imaged volumes of brain tissue must be densely labeled. This means that the cell type must be known for each voxel in the imaged brain tissue. Manually

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

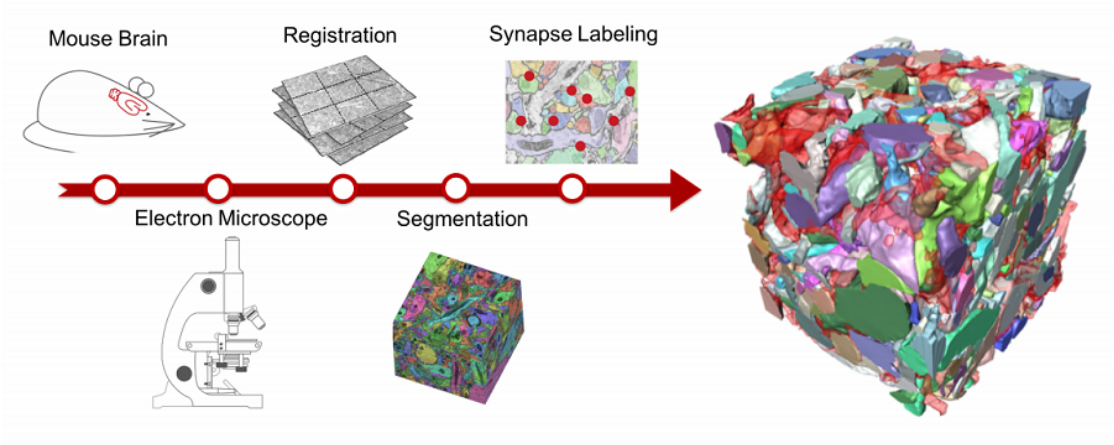


Figure 3.2: **Steps of the connectomics workflow.** Source: [Moh17]

labeling such large volumes is almost impossible. Therefore, scientists use machine learning-based segmentation algorithms to densley label the data.

Now, we describe each step of the connectomics workflow in detail. Figure 3.2 illustrates the connectomics workflow.

3.2.1 Data Acquisition

Our collaborators start data acquisition by dissecting solid blocks of brain tissue specifically stained for electron microscopy. Next, they image individual slices with an electron microscope at a pixel resolution of 5×5 nanometers [KMWL08], with a slice thickness of 15 nanometers. This resolution requires around 450 imaged slices of brain tissue to image a volume of 5 cubic microns. Conventional EM imaging of volumes of brain tissue requires the cumbersome manual work of skilled experts. Newer methods, however, have automated the slicing process and, therefore, allow to image larger volumes of brain tissue. An example of an imaged slice of brain tissue is shown in Fig. 3.3. Contours of individual nerve cells are visible. Darker image regions within nerve cells correspond to mitochondria.

3.2.2 Image Registration

Image registration describes the process of transforming different images into the same coordinate space. To align all slices with each other, various registration methods can be used. Thereby registration approaches are categorized into three groups based on the applied image transformations [Rai19]. *Rigid transformation* use image translation and rotation to align image pairs with each other. *Affine transformations* extend rigid transformations by allowing rescaling images to match them with other images. *Deformable/Curved transformations* also allow image distortions. Independent of the applied transformations, robust image registration requires a similarity measure to

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

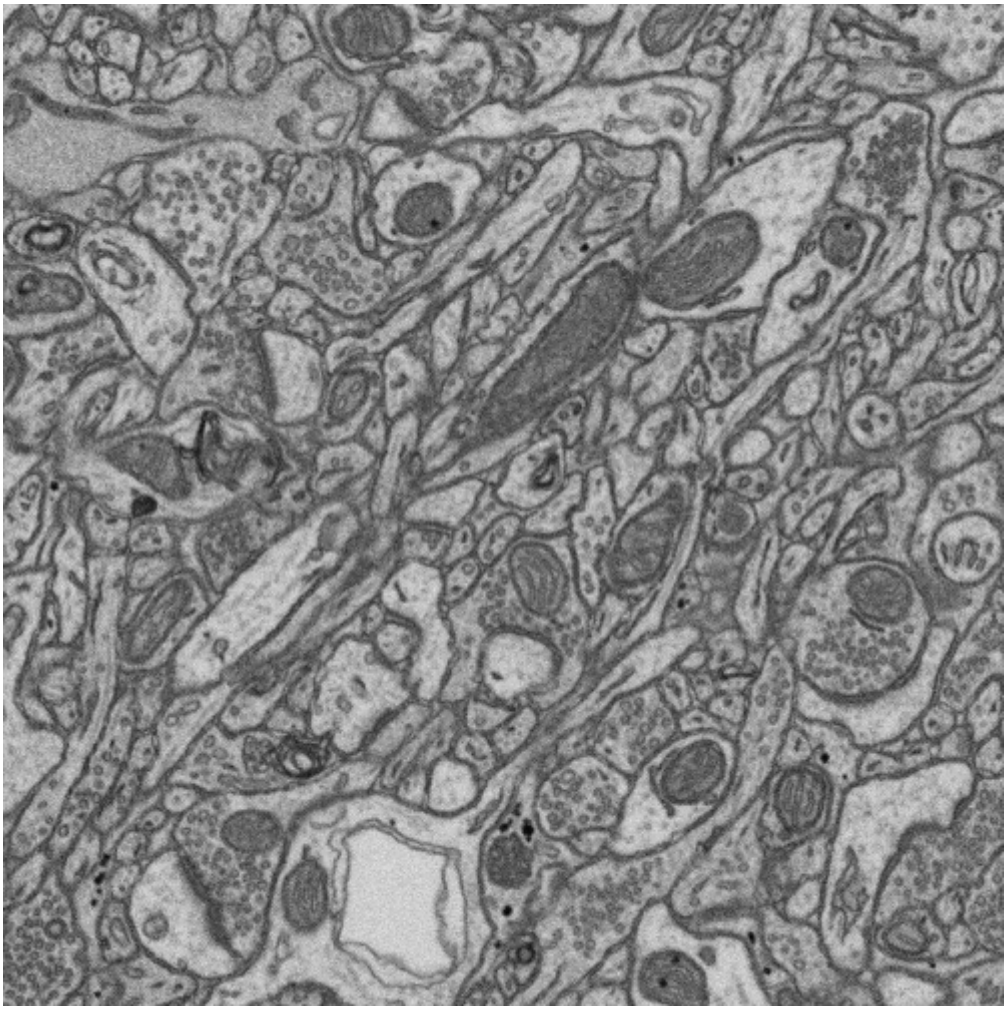


Figure 3.3: **Slice of brain tissue imaged with electron microscopy** at a resolution of 5×5 microns. Data provided by Graham Knott [KMWL08].

evaluate the quality of image alignments. Several similarity methods have been proposed, including image-intensity and geometry-based methods [Rai19]. Additionally, non-rigid image transformations need to use interpolation between pixels to fill the new space.

3.2.3 Segmentation & Reconstruction

Image segmentation is an active area of research. Here, the goal is to assign a semantic label to each pixel of an image. For example, we want to know for each pixel in an EM slice image, which neuronal structure it represents (see Fig. 3.4). The scientists' goal is to densely label the data, meaning that they segment every cellular structure in the given tissue block. Dense labeling is essential for spatial neighborhood analysis because it allows us to look at all structures and cellular substructures within a local neighborhood.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

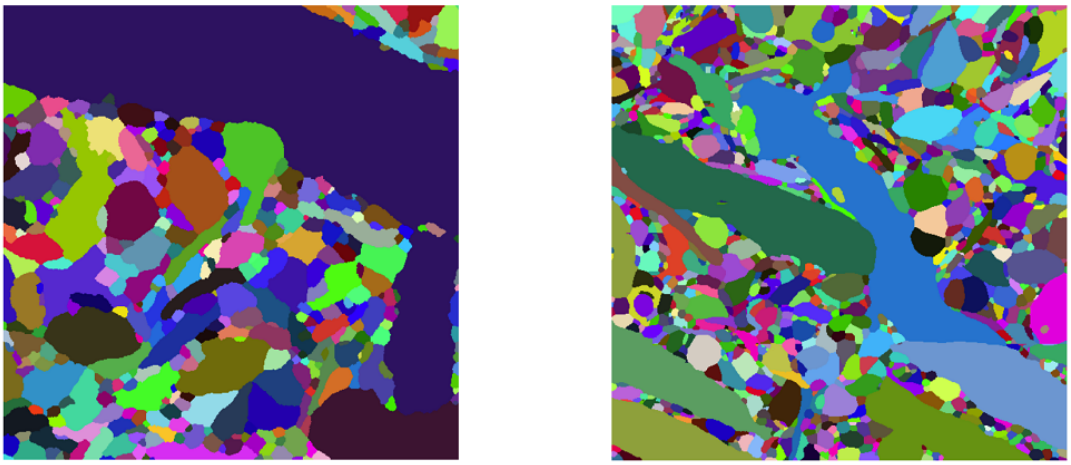


Figure 3.4: **Segmentation masks of EM connectomics data.** Each color relates to a neuronal structure. Source: Brian Matejek [MHL+17]

Conventional methods for image segmentation include the semi-automatic region-growing approach or the automated watershed image segmentation. However, most methods used for instance segmentation in EM connectomics data, are based on deep learning models. For instance, Gonda et al. [GWP21] use consistent recurrent neural networks to create binary segmentation masks for EM images with spatiotemporal consistency. Lin et al. [LW19] provide *PyTorch Connectomics*, a toolbox for instance segmentation of EM connectomics data. They integrate automatic and semi-automatic approaches as well as segmentation techniques for cell organelles such as mitochondria [WLFB+20]. Furthermore, scientists use a wide range of other tools to detect and segment neurites and synapses using semi-automatic approaches like Ilastik [SSKH11], TrackEM [CSS+12], or VAST [BSL18]. After the segmentation process is completed, 3D surface meshes are extracted from the segmentation stacks. We use these mesh representations to render neuronal structures in NeuroKit.

3.2.4 Visualization

In the last step, scientists want to create informative visualizations giving them detailed insights into the data. Here, the goal is to transform the data into actual knowledge that helps scientists to understand how the brain functions. For instance, NeuroKit uses surface meshes to quantify distances between neuronal structures. These distances are then visualized using custom charts and diagrams. This information helps our collaborators to characterize brain tissue more precisely and form new hypotheses on how neuronal structures interact with each other. The biggest challenge in visualization for connectomics is to handle the vast amounts of data efficiently that the previous steps of the pipeline produce. First, we face computational challenges like loading large data

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

sets of high-resolution brain tissue into memory. This requires sophisticated rendering techniques like virtual memory approaches [HBJP12] or efficient empty space skipping in volume rendering [HAAB+17]. On the other hand, visualization must avoid a cognitive overload of the user when analyzing complex data sets. Therefore, suitable complexity reduction methods, such as data abstractions [AABS+14, Moh17], must be invented. In addition to extracting domain knowledge from the data, visualization approaches are also used to guide the proofreading process of the previous *Segmentation & Reconstruction* step. Deep learning-based models often produce errors when segmenting instances of neuronal structures. Here, visualization approaches help to identify and correct errors quickly [AABH+16, GWB+21].

3.3 Data

In NeuroKit, we display the reconstructed structures as surface meshes and also visualize the original EM-image volume. Our collaborators have currently collected and reconstructed six *focused ion beam scanning electron microscope* (FIB-SEM) data sets, which are all five cubic microns in size [KMWL08]. Three data sets show brain tissue of four-month-old mice, while the other three show tissue of 24-month-old mice [CWB+18]. For all six data sets, dendrites, axons, synapses, mitochondria, boutons, and spines have been reconstructed. All six data sets are taken from layer one of the somatosensory cortex. Figure 3.5 gives an overview of all six data sets. The left column shows a 3D rendering of the respective data set and connectivity information as a node-link diagram. Having data from specimen of different age groups available is helpful to understand how the mammalian brain develops and changes during the aging process.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

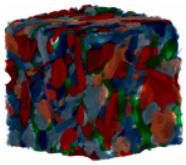
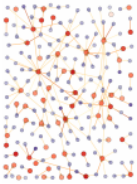
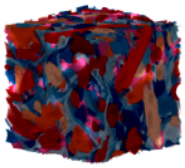
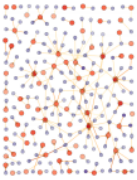
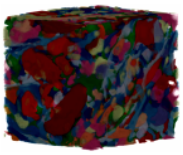
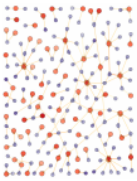
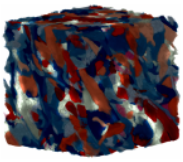
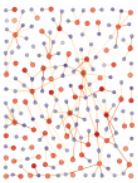
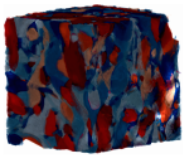
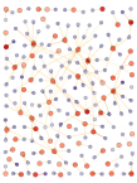
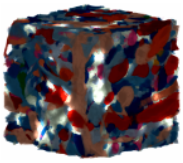
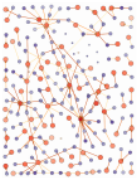
Dataset		Description	Classification
		Mouse cortex 1 4 months 831 x 831 x 280 5μm x 5 μm x 5μm	Synapses: 189 Spines: 79 Dendrites: 66 Boutons: 115 Axons: 274
		Mouse cortex 2 24 months 833 x 833 x 468 5 μm x 5 μm x 5 μm	Synapses: 212 Spines: 104 Dendrites: 84 Boutons: 102 Axons: 381
		Mouse cortex 3 4 months 999 x 999 x 449 5 μm x 5 μm x 5 μm	Synapses: 167 Spines: 152 Dendrites: 81 Boutons: 237 Axons: 274
		Mouse cortex 4 24 months 1001 x 1001 x 304 5 μm x 5 μm x 5 μm	Synapses: 185 Spines: 81 Dendrites: 87 Boutons: 94 Axons: 285
		Mouse cortex 5 24 months 1001 x 1001 x 464 5 μm x 5 μm x 5 μm	Synapses: 171 Spines: 103 Dendrites: 74 Boutons: 82 Axons: 355
		Mouse cortex 6 4 months 998 x 998 x 268 5 μm x 5 μm x 5 μm	Synapses: 211 Spines: 135 Dendrites: 85 Boutons: 102 Axons: 286

Figure 3.5: **Six different data sets acquired by our collaborators [KMWL08].** The first column shows a 3D rendering of the respective data set and the connectivity of nerve cells as a node-link diagram. The data has a cubic geometry since it shows a small subvolume of the mouse brain. The second column shows information data of the data set and the third column holds information about the segmented structures. Source: [ACA+19]

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 4

Goal & Task Analysis

Following a problem-driven design study approach by Sedlmair et al. [SMM12], we have identified domain goals, associated analysis tasks, and general system requirements in semi-structured interviews with four experienced neuroscientists at the University of Turin and the Harvard Center of Brain Science. All scientists work with 3DEM data and have more than ten years of experience in the field. Over several months, we regularly met with our main neuroscience collaborator to demonstrate design prototypes and collect feedback.

4.1 Design Study Methodology

Before describing the results of our design study in detail, we provide an overview of the design study approach [SMM12] we have followed. Figure 4.1 illustrates the different steps of the approach. The approach is divided into the three phases *precondition*, *core*, and *analysis*. Each of these phases is again split into several subparts.

The *precondition phase* aims to plan the research project by learning about the state of the art in the scientific field. This helps to make informed decisions on the goals in the design study. Also, potential collaboration with domain experts should be assessed in this phase. Here, it is crucial to check whether the domain-specific problem is suitable for a visualization design study. Visualization researchers should evaluate if the problem of the potential collaborators requires research or just engineering efforts to solve. Other aspects include the availability of data and time constraints. After these considerations, roles in the team of researchers must be distributed. Roles range from the *front-line analyst*, who will use the resulting visualization approaches, to *gatekeepers*, responsible for approving or declining feature requests and other leadership tasks.

In the *core phase*, visualization researchers and domain experts must formulate a detailed problem statement. After that, the visualization design, data abstractions, and visual

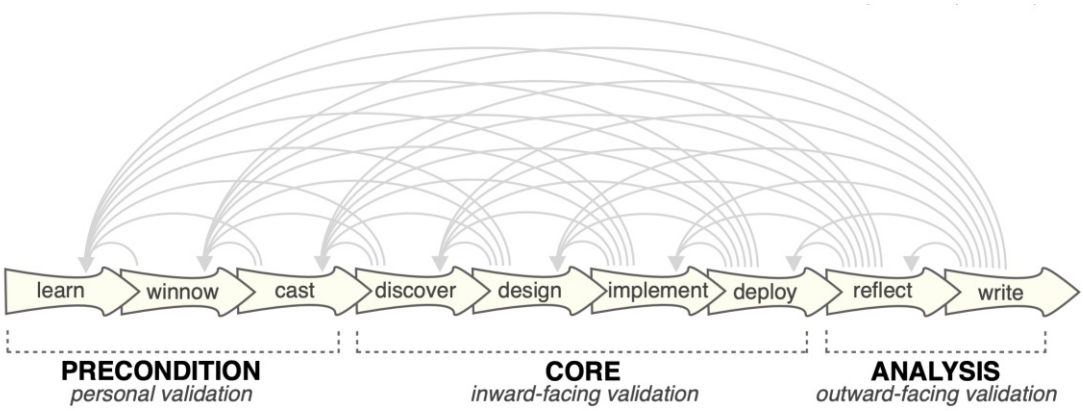


Figure 4.1: Overview of the design study methodology approach by Sedlmair et al. [SMM12]

encodings are specified. In the next step, the implementation of software prototypes takes place. Here, it is crucial to follow a rapid prototyping approach delivering fast results. Spending too much time on software development by aiming to create perfect solutions is a common pitfall in this phase. Once the implementation phase is over, the project and the code should be open-sourced, and the effectiveness of the approach must be evaluated by collecting feedback.

Finally, in the *analysis phase*, the focus is on reflecting on the process and then writing down the design study results. Evaluating the team performance as well as the results of the study are part of the reflection. Paper writing requires a sufficient amount of time and must be planned accordingly. Often, the writing phase requires reconsiderations of the data abstraction to articulate them accordingly.

Figure 4.1 sketches the design study approach as a linear process. However, phases overlap, and the backward arrows indicate iterations between the single phases. Now, we report on the results of our discussions with domain experts by outlining domain goals, analysis tasks, and requirements.

4.2 Domain Goals

Our collaborators' main objective is the detailed analysis of local neighborhoods in their neuronal 3DEM data sets. In particular, they want to analyze, compare, and correlate morphological features of interest in selected cells, such as spines, boutons, or mitochondria, to discover previously unknown biological patterns. The particular domain-specific goals of the neuroscientists are:

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

G1 - Proximity Analysis for Subcellular Structures

The collaborators are concerned with the spatial aspects of how cellular substructures relate to each other. On the one hand, they are interested in analyzing **distances** between a varying number of structures (e.g., *What are the distances of the closest three synapses to a mitochondrion, and what is their average size?*). On the other hand, they are interested in **distributions** (e.g., *What percentage of a mitochondrion is closer than 1 nm to the cell boundary?*). Our collaborators want to investigate whether patterns indicate connectivity strength based on the underlying spatial arrangement of cell substructures.

G2 - Comparative Analysis of Neuronal Structures

In addition to looking at a single spatial neighborhood, our collaborators need to compare neighborhoods and structures of interest. Their goal is to perform comparisons with **different biological targets**, to compare between: a) different data sets (e.g., specimen from different developmental stages), b) different instances of a structure of interest (e.g., dendrites), or c) different attributes of a single structure of interest (e.g., synapse distances to a single mitochondrion). Additionally, they need to be able to perform comparisons of **different cardinalities**, ranging from a detailed comparison of two structures with each other to comparing hundreds of structures at the same time. This feature allows them to test whether local findings also apply to a larger data set and, conversely, allow them to drill down from a larger perspective to highly detailed local instances.

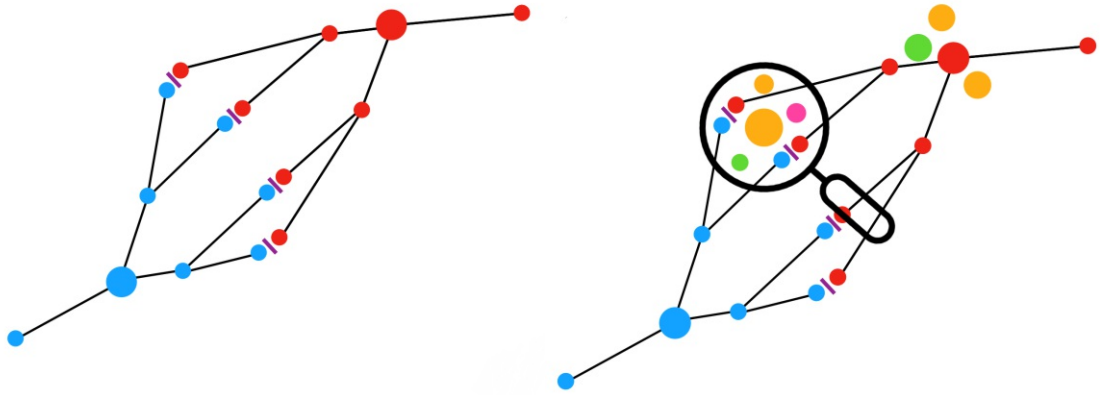
G3 - Intuitive Analysis of User-Specific Hypotheses

Our scientists want to be able to fine-tune their visual analysis to their current scientific objective. In addition to general comparisons and proximity analysis in their data, they want to customize analyses steps and define which visualizations to use for the domain-specific tasks. For example, one of our collaborators is currently focusing on the analysis of mitochondria in dendrites. We describe the specific tasks for this analysis scenario in detail in Chapter 8.

4.3 Analysis Tasks

Based on the above goals, we have derived a set of five analysis tasks that need to be supported in NeuroKit to aid scientists in the analysis of spatial neighborhoods in their data. Note that we distinguish between connectivity analysis and spatial neighborhood analysis. Figure 4.2 illustrates the difference between the two analysis tasks. In the following paragraphs, we explain these analysis tasks and also link them to the NeuroKit domain goals:

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) Connectivity analysis (b) Neighborhood analysis

Figure 4.2: **Connectivity vs. neighborhood analysis.** Connectivity analysis of neuronal structures (a) compared to spatial neighborhood analysis (b). Our collaborators are specifically interested in the latter.

T1 - Explore original EM image data and segmented structures

Before diving into a detailed analysis, scientists first need to get an overview and verify the correctness of the segmentation (**G1, G3**). Verifying data segmentation is essential to ensure a solid analysis phase and check whether any mistakes have been overlooked in the proofreading phase.

T2 - Explore the spatial neighborhood of a single structure

Once a structure of interest has been identified, domain scientists explore spatial features and morphology in more detail (**G1, G3**). Thereby, the spatial features of the neighborhood and the structure's 3D shape are of specific interest. Spatial features and morphologies vary highly between structures of interest.

T3 - Quantify distances and distributions in a spatial neighborhood

Domain experts are interested in quantifying the proximity and spatial arrangement of structures in a spatial neighborhood (**G1, G3**). Quantifying the neighborhoods facilitates a more detailed analysis of neighborhood features and, therefore, allows more accurate observations.

T4 - Compare the spatial neighborhoods of several structures

Analyzing the variance of certain features in different local spatial neighborhoods is essential for discovering trends and patterns when studying biological features (**G2, G3**).

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

This analysis task is closely linked to analysis task T3. Quantifying neighborhood features leads to improved comparison possibilities.

T5 - Compare neighborhoods of different data sets

Scientists want to compare different specimens to analyze the stability of certain neighborhood features and differences between them (**G2, G3**). For instance, our collaborators are interested in comparing brain tissue of young and old mice to analyze how the aging process affects the development of neuronal structures.

4.4 Requirements

To support the goals and tasks not just for one specific expert user but also for many different users and their individual domain questions, we have extracted some design requirements that NeuroKit must fulfill.

R1 - Flexibility and Expandability

While all of our collaborators want to analyze spatial neighborhoods and compare structures of interest, the specific details vary for each scientist. For example, one of the collaborators is mainly interested in mitochondria, while another scientist focuses on endoplasmic reticula. To support both (and other) usage scenarios in NeuroKit, the system has to be customizable and expandable to allow scientists or technicians to adjust the provided visualization capabilities for specific domain questions (**G3**).

R2 - Scalable Comparison

To support the visual comparison of different biological targets and at different cardinalities, our system needs to provide comparisons at different scales and allow users to switch between them easily (**G2**). This is necessary to identify trends in larger sets of neighborhoods but also allows a detailed investigation of small multiples.

R3 - Ease of Use

In addition to being flexible and expandable, the system must also support neuroscientists with little or no programming experience. Biological data, and especially neuronal data, is inherently complex. Therefore, we need to limit the cognitive load for the user by employing methods that automatically guide and support them in the exploration and analysis process. Hence, NeuroKit should provide a human-readable interface and file format for customization and automatically guide the user in their specific analysis tasks (**G1-3**).

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

CHAPTER 5

NeuroKit Toolkit Design

We have designed NeuroKit with two main considerations in mind. First, during discussions with our collaborators, we realized that they want to shift their analysis from being centered on brain connectivity to a more detailed analysis of spatial neighborhoods. These neighborhoods in EM data can be extremely complex and can include dozens of cellular structures and substructures. Second, we want to create a highly flexible and customizable toolkit that can support a broad range of specific domain tasks and guides users in their analysis steps by automatically suggesting appropriate visualizations, view parameters, and interaction modes. Neuroscientists often have particular analysis goals that differ even between scientists in the same research group. Visualization tools that are designed to be general enough to support different types of analyses often result in overwhelming user interfaces and unintuitive workflows. Alternatively, tools designed for a specific analysis task typically do not generalize well to other tasks. NeuroKit aims to combine the best of these two worlds by providing a general framework that is expressive and powerful but can be customized to support individualized workflows. By allowing scientists to customize their analysis workflow to best suit their needs, we minimize visual complexity, streamline the user interface, and simplify user interaction.

In this chapter, we focus on designing the configurable toolkit and framework aspects of NeuroKit. The visual neighborhood analysis and comparison approach is then described in Chapter 6. Fig. 1.2 shows the overall design of NeuroKit. Based on the custom user-defined analysis scenario, NeuroKit guides users in their analysis by suggesting fitting visualizations, adjusting view parameters, and offering different comparison modes.

5.1 Custom Analysis Scenarios

NeuroKit allows users to define custom analysis scenarios that fine-tune the general neighborhood analysis capabilities of our framework to a highly customized workflow (**R1**).

5.1.1 Scenario Specification

Conceptually, a domain-specific analysis scenario defines the analysis a neuroscientist wants to perform and breaks it down into individual and concrete domain tasks and elements. These domain tasks are defined in terms of data mappings, default visualizations, comparison modes, and supported user interactions. We will refer to the tasks in a user-defined analysis scenario as *domain tasks* to distinguish them from the general analysis tasks identified in our goal and task analysis in Chapter 4. Below, we describe the concrete elements comprising a scenario.

Scenario:

An analysis scenario represents a higher-level question a domain scientist might have, such as "*How are mitochondria arranged in dendrites?*", "*How do synapses cluster around mitochondria*", or "*How do mitochondria cluster around synapses?*".

Domain tasks:

Each analysis scenario comprises one or several domain tasks, which are specific analysis steps a user needs to perform to answer the higher-level scenario question. For example, domain tasks in the mitochondria scenario above would be *analyzing the synaptic neighborhood of a single mitochondrion*, or *analyzing the distance from a mitochondrion to the cell boundary*, or *analyzing the placement of mitochondria within a nerve cell*. To support these steps, each domain task can define visualizations, comparison modes, and interaction parameters that help the user in the analysis.

Target structure:

Each scenario can define a target structure (TS), which is the main biological structure under investigation. Target structures can be cell organelles like mitochondria or other cellular structures such as dendrites or synapses. Specifying a target structure ensures that visualizations and user interface elements defined in specific domain tasks focus on those structures (e.g., center their view on them, or show a list of all instances of that structure in the data set). Note that we distinguish between the general biological structure (i.e., target structure) and specific instances of that structure a user selects during their analysis (i.e., structures of interest).

Visualizations and data mappings:

Each domain task contains a set of customized visualizations to show those data elements that are best suited to complete the specific task. For example, to see the distance between a mitochondrion and the cell boundary, we can define a visualization that color codes the distance onto the mitochondrion's surface in a 3D rendering or show a violin plot that encodes how the closest distance varies over the surface of a mitochondrion. NeuroKit supports a set of standard information visualization views, such as bar charts,

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

scatterplots, violin plots, or cumulative histograms, as well as several custom views for neuroscience tasks (see Chapter 8). Each view can be customized by changing the specific data mapping, such as which attributes to show on the axes of a scatterplot. Furthermore, NeuroKit is extensible and allows to implement and plug in additional views that are not part of the original framework (see Chapter 7). Finally, we can define comparative views for different cardinalities. This means that we can specify a high-detailed view for low-cardinality comparisons and overview visualizations for high-cardinality comparisons (**R2**).

Interaction parameters:

In addition to specifying visualizations and tasks, a scenario can also define interaction parameters, such as user-adjustable sliders, to filter data before visualization. For instance, NeuroKit can be configured such that the number of bins in a cumulative histogram or specific threshold distances are adjustable by a slider.

5.1.2 Analysis Definition File

Custom analysis scenarios are defined by users in the form of *analysis definition files* (ADFs). ADFs are written in a human-readable JSON format and do not require any programming skills (**R3**). Fig. 1.2 (a) shows the structure of an ADF and Listing 7.10 and 7.13 give concrete examples. NeuroKit reads the ADF upon startup and configures the user interface accordingly. We provide sample ADFs that the users can further fine-tune. The NeuroKit documentation [Tro21] includes detailed information about supported visualizations and their parameters.

5.2 Automatic User Guidance

Based on the active analysis scenario, NeuroKit automatically adjusts the user interface, views, and interactions and helps guide users through the analysis (**R3**).

5.2.1 User Workflow

The user workflow starts by loading a custom analysis scenario and data (see Fig. 5.1a). First, users perform an initial exploration in 3D and select structures of interest for further investigation. Next, users can perform the custom domain tasks of their loaded scenario, such as comparing the selected structures in detail to extract distances, detect distribution patterns or test hypotheses. In all these steps, users are supported and guided by NeuroKit to streamline the analysis process and minimize the mental load (see Fig. 5.1e). Fig. 5.1 shows the user workflow and which parts of the framework are adjusted automatically. After completing the analysis of one domain task, users can switch to another domain task, refine the data selection, and adjust the scale of the comparison method iteratively.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

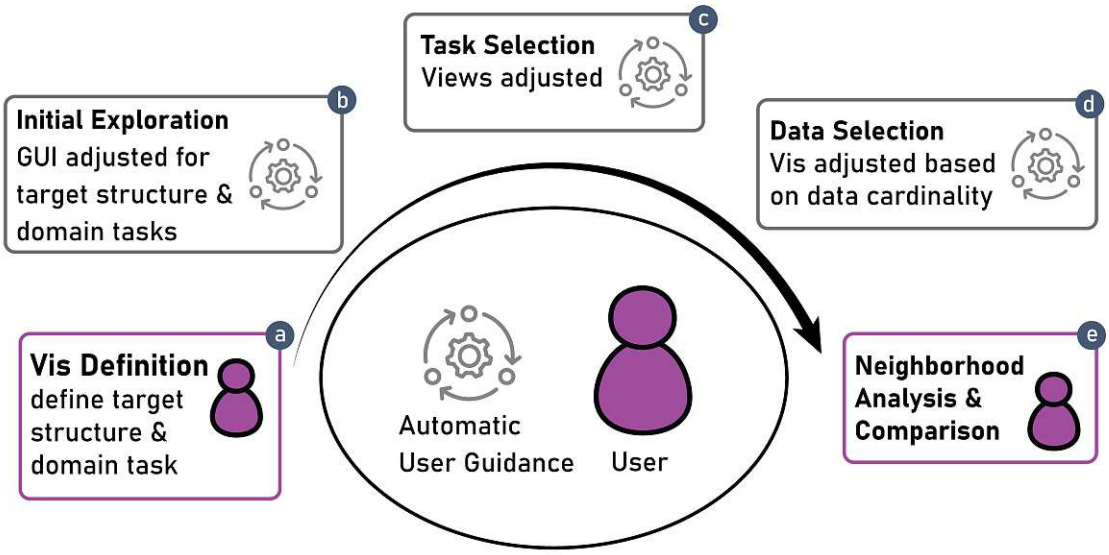


Figure 5.1: **NeuroKit workflow and automatic user guidance.** NeuroKit guides and supports users throughout the analysis workflow to minimize mental load and support specific analysis tasks. We show user actions in purple, and automatic guidance actions in gray. A user can customize the visual analysis by specifying a user-defined analysis scenario. Based on the analysis scenario, we automatically adjust the visual user interface and propose the most appropriate visualization and comparison parameters.

5.2.2 Guidance and Automatic Adjustments.

NeuroKit includes automatic guidance features in each step of a scenario analysis session.

Initial Exploration

(Fig. 5.1b): Whenever a scenario defines biological target structures (e.g., mitochondria), we automatically highlight those structures throughout the system. We create a list view of all instances of that structure (e.g., all mitochondria in the data set) and blend out other structures in the list view. The target structure also constitutes the center of a spatial neighborhood. Any spatial neighborhood analysis will be centered around that structure.

Domain Task Selection

(Fig. 5.1c): During the exploration step, depending on the user-defined scenario and selected domain task, NeuroKit automatically updates the list of available visualizations and adjusts view parameters. For example, depending on the selected domain task and its settings, we adjust which cell organelles are visible by default in the 3D view of a region of interest.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Data Selection

(Fig. 5.1d): For each domain task, NeuroKit can automatically adjust the level of detail in our scalable comparison framework (see Chapter 6), depending on the number of structures selected for comparison. This allows us to adjust the level of detail in a comparative visualization based on data cardinality.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 6

Visual Neighborhood Analysis

Analyzing and comparing spatial neighborhoods is challenging, especially in biological tissue that is highly variable and exhibits a complex three-dimensional structure. Looking at not just one but multiple neighborhoods solely in 3D easily leads to cognitive overload and does not allow quantitative comparisons. Therefore, NeuroKit offers two complementary visualization methods. First, we can show juxtaposed high-detail 3D renderings of the selected local neighborhoods. This allows neuroscientists to explore the original 3D data and get an intuition about the morphology and spatial arrangement of structures of interest. In addition, we show linked *analysis views* that support a more quantitative analysis and comparison of neighborhoods. Fig. 6.1 shows the entire user interface of NeuroKit.

6.1 Scalable Comparison Framework

Conceptually, NeuroKit supports two different levels of comparison, which are orthogonal to each other. On the one hand, we support different biological targets of comparisons, where targets relate to biological specimens, structures, or cellular substructures. On the other hand, we support comparisons of different cardinalities (i.e., a varying number of objects in a comparison). Both approaches allow comparisons to range from overview to detail in both biological and visualization terms (**R2**). The scalable comparisons are fully integrated into the custom analysis scenarios framework, allowing NeuroKit to automatically suggest and adjust comparative views based on the current domain task and the number of selected structures.

Comparison Targets

NeuroKit supports three different comparison targets. First and foremost, it supports comparing different instances of the same biological target structure (e.g., comparing

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

two dendrites). Second, it supports the comparison of cellular substructures of the same structure of interest (e.g., synapses of the same dendrite). Third, NeuroKit supports the comparison of different data sets (e.g., data sets of neuronal tissue in different developmental stages). Users can define comparison targets in the ADF by setting the *target structure* and specifying the data mapping for comparisons. NeuroKit then automatically adjusts the user-interface (UI) and views based on the specified comparison target.

Cardinality of Comparisons

In their analysis, scientists often move from a single interesting structure to many to verify whether identified features correspond to larger trends or patterns. Alternatively, they usually start with an overview of many structures and then drill down to see more details. Therefore, we designed a scalable comparison framework to support comparisons of sets with different cardinalities. In particular, NeuroKit provides comparisons at three different scales (Fig. 1.2). The *low-cardinality* comparison mode is designed to compare a handful of objects (two to five) and offers high-detail juxtaposed views on the objects being compared, typically arranged as small multiples. A highly detailed view takes up much screen space and therefore only supports a low number of objects that can be compared. The *medium-cardinality* comparison mode supports up to a dozen structures, while the *high-cardinality* comparison view scales up to hundreds of objects. The latter view is the most compact visualization and makes larger trends, patterns, and outliers visible.

NeuroKit can either automatically switch between comparison modes, depending on the number of selected objects, or let users choose their desired comparison mode. The actual visualization for each comparison mode is based on the domain tasks specified in the analysis definition file. Figures 8.2, 8.4, and 8.6 show the different cardinality views for domain tasks in the "Mitochondria Analysis" usage scenario (Chapter 8).

6.2 Visual Elements

In this Section, we describe the different visual elements of our application (see Fig. 6.1) and how they support visual neighborhood analysis. In addition, we outline how our design choices and interaction techniques amplify the understanding of our data.

Scenario Specific User Interface

NeuroKit automatically parses the ADF and creates UI elements specified for a user-defined analysis scenario. For example, for analyzing mitochondria, scientists might want a slider to interactively change the distance threshold of what they consider a "close" synapse.

Furthermore, we automatically show icons for quick previews of the three different comparison levels for each domain task in the scenario (see Fig. 6.1e). The icons allow

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

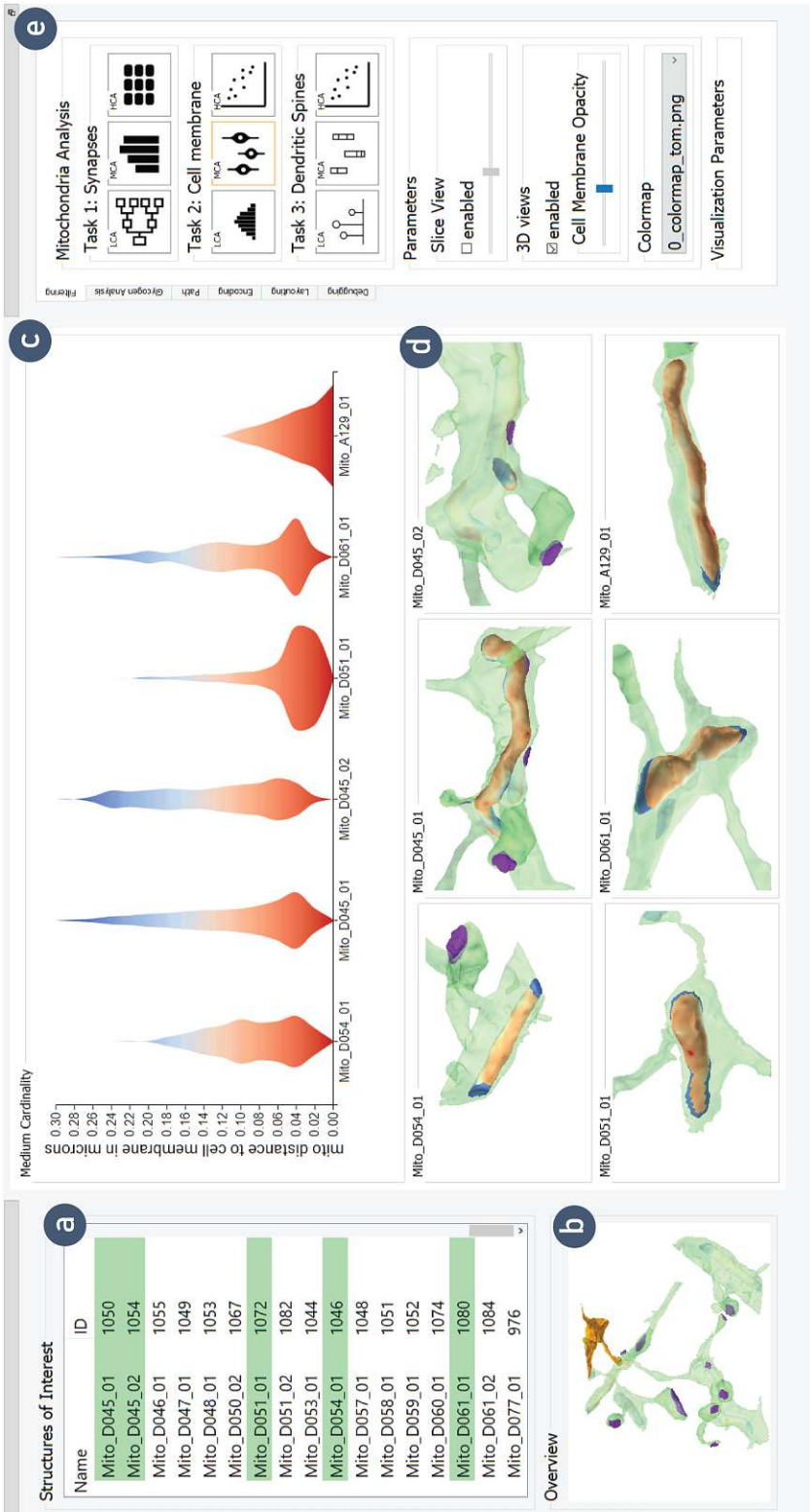


Figure 6.1: **NeuroKit user interface.** (a) List view for selecting structures of interest. (b) 3D overview of all selected structures of interest. (c) Analysis view for detailed comparisons of the selected local neighborhoods. (d) Detailed 3D neighborhood views. (e) Panel for user-defined analysis scenario, showing suggested comparison modes and custom UI elements.

users to choose their current domain task quickly and to switch the comparison mode manually, and it gives a preview of the individual comparative visualizations.

3D Neighborhood View

The neighborhood view (see Fig. 6.2) allows the detailed examination of the local spatial neighborhood of a user-selected structure of interest (**T2**). This helps neuroscientists to analyze the morphology and three-dimensional anatomy of the selected structure while at the same time seeing nearby cellular substructures. For example, in the neighborhood view, we can show a selected mitochondrion, its respective axon or dendrite, as well as nearby synapses. We display the selected structure of interest and the surrounding region in a high-resolution 3D surface rendering. In addition, we may also display a slice view of the original EM data that can be moved through the neighborhood view (**T1**). This feature is essential for neuroscientists for two reasons. First, it allows them to proofread and check segmentation results on-the-fly. Therefore, the 3D renderings can also be disabled to avoid occlusions when verifying the segmentations. Additionally, it allows them to discover and analyze features in the EM data that have not been segmented and reconstructed yet, such as individual glycogen granules and their distribution.

We have implemented the neighborhood view as small multiples (see Fig. 6.1d). We juxtapose views of multiple local neighborhoods next to each other (**T4, T5**). Small multiples can be linked to each other for synchronized navigation (i.e., zooming, rotation, EM slice view placement) between all views.

Analysis View

While the 3D neighborhood views are great for getting a spatial understanding of a neighborhood, it is still hard to compare multiple neighborhoods more quantitatively. Therefore, NeuroKit offers *analysis views* (see Fig. 6.1c) that abstract certain neighborhood characteristics and support direct comparisons between neighborhoods (**T3, T4, T5**).

The actual visualization shown in the analysis view is based on the number of selected structures (i.e., the cardinality of the current comparison) and the specific analysis scenario (i.e., the settings in the analysis definition file). Analysis views are connected to the 3D neighborhood views by brushing and linking to allow users to easily connect the content of the analysis views to the neighborhood views. After spotting an interesting neighborhood feature in the analysis view, the user can study this feature in greater detail in the 3D representation of the neighborhood.

Spatial Context View

To provide an initial overview of the entire data set and all its segmented structures, NeuroKit provides a spatial context view (see Fig. 6.1b). With this view, scientists see the spatial location of selected structures in the context of the entire volume. They can

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

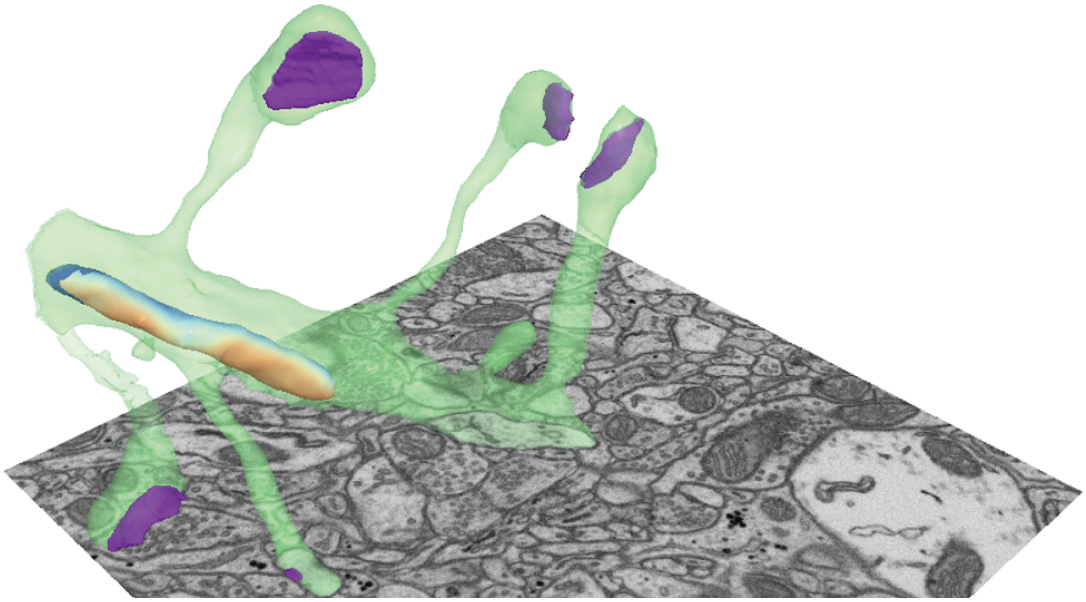


Figure 6.2: **Detailed three-dimensional neighborhood view.** The dendrite (green) developed spines that form synapses (purple), and an enclosed mitochondrion (surface color-coded structure). Orange/reddish color on the surface indicates close proximity between the mitochondrion and the dendrite's cell membrane, while blue denotes a greater distance. The 3D rendering is overlayed with a 2D slice of the original EM data.

examine how different local neighborhoods are related to each other (**T1**). For instance, it is interesting to our collaborators whether specific neighborhoods overlap and share synapses. Furthermore, it is possible to detect clusters in the spatial overview. For example, mitochondria clusters could indicate a region of high metabolic activity. We link the spatial overview to the individual neighborhood views. When hovering over a structure in the neighborhood view, we highlight the respective structure in the spatial context view and vice versa.

Selecting Structures of Interest

To easily select structures of interest, we provide a list view of all segmented objects in the data set (see Fig. 6.1a). Suppose a biological target structure has been defined in the analysis definition file. In this case, we can automatically filter out all other structures and only show segmented objects of the target type (e.g., dendrites). This facilitates a fast selection of structures of interest without time-consuming search through all segmented structures.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

6.3 Interaction

We have chosen specific interaction techniques to facilitate fast and convenient visual analyses of our data. Navigating view settings in multiple neighborhood views can be cumbersome. That is why users can optionally link zooming for all neighborhood views. Having both linked and individual zoom allows users to quickly view all neighborhoods from the same distance while still enabling a detailed and easy examination of unique neighborhoods. Furthermore, we support brushing and linking between analysis views, neighborhood views, and spatial context views. Hovering over a specific region in the analysis view highlights the entire neighborhood view it corresponds to. This is especially helpful when comparing many structures to easily spot the corresponding neighborhood view. We also highlight individual structures like synapses or dendritic spines in the 3D views once selected in the analysis view. Thus, users can easily connect the insight they gained from the analysis view to the detailed 3D representation.

We support two different interaction modes, depending on the cardinality of the current comparison. For *low-scale comparisons*, users start by choosing specific structures from the list view, which are then shown as small multiple neighborhood views, and subsequently abstracted and quantified in the analysis view. For *high-scale comparisons*, on the other hand, we automatically display all instances of the biological target structure in the analysis view. In the next step, the user can select individual structures of interest in the analysis view, which are then shown in the detailed neighborhood views.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 7

Techniques

In this chapter, we describe the NeuroKit architecture, all required data preprocessing steps, techniques used for application development, and a detailed explanation of the structure of ADFs.

7.1 NeuroKit Architecture

We have implemented NeuroKit as a modular application to split up *NeuroKit-core* from the *user interface* and the *visualization stack*. The *visualization stack* provides a general API for all visualizations in our system. Whenever a new analysis scenario is loaded, the core module re-initializes and updates the user interface and initializes the different views of the visualization stack specified in the analysis definition file (ADF). In case a usage scenario requires a visualization that has not been implemented in NeuroKit yet, the visualization stack is extensible. Additional D3-based data visualization modules have to implement our visualization interface and can then be further customized via the analysis definition file. Since NeuroKit already supports standard visualizations (e.g., scatterplots, bar charts, etc.), this feature is primarily meant to support novel custom visualizations specific to a particular usage scenario.

7.2 Data Preprocessing

In NeuroKit, we display surface meshes of the reconstructed structures and the original EM image volume. We extract the mesh data from segmentation stacks using Neuromorph [JNC+15] and compute the closest distances between all structures of interest using the CGAL library [FGK+00]. Our collaborators provided other mesh attributes such as spine lengths, spine volume, or surface area. We have defined a standardized file format to read arbitrary data sets into NeuroKit. Mesh data must be specified in the

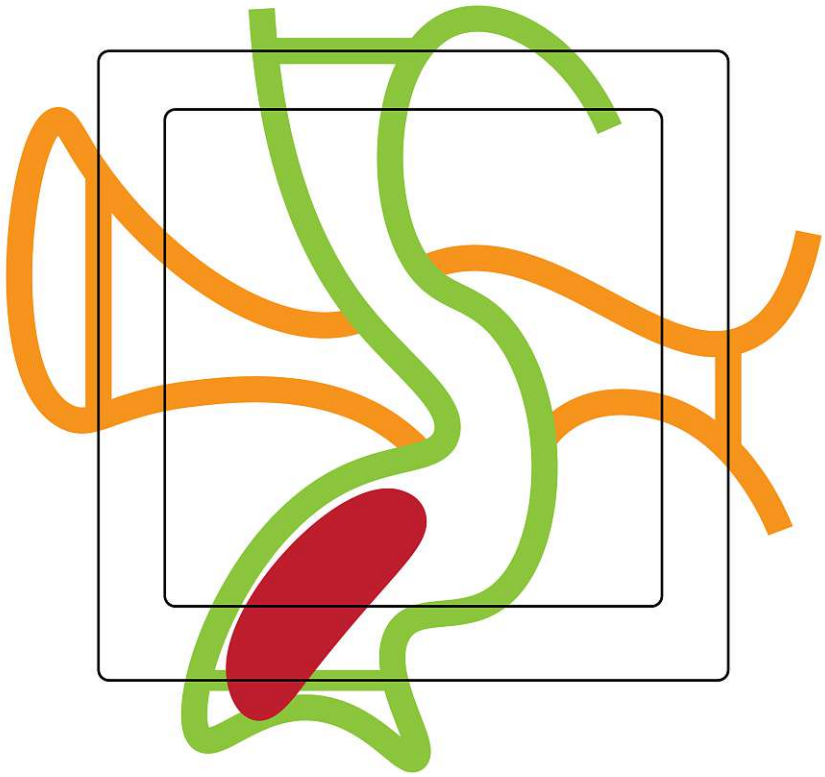


Figure 7.1: **Volume boundary handling for distance computations.** Neurites (orange, green) and cell organelles such as mitochondria (red) are often cut off by the volumes cutting plane (outer box). To prevent a distance distribution from being incorrectly calculated, we introduce a margin (space between boxes) that is not considered for distance distributions.

wavefront data format, and we use a standardized naming convention for all objects to make data loading easier.

7.2.1 Distance Computation

We use a tree data structure to perform distance computations [ATW20] between 3D objects. We construct a hierarchy of axis-aligned bounding boxes (AABB-tree) from a set of triangles. The distance between a point and the object can then be computed by determining which bounding box contains the point. In NeuroKit, we perform two types of distance computations. First, distances between objects are calculated to determine the proximity between different structures. Second, we compute distance distributions of cell organelle surfaces to their respective cell membrane to investigate the placement of cell organelles.

We often only view fractions of neurites and cell organelles since they are truncated by

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

the volumes' boundary. At the intersections of the neuronal structures with the volume boundary, new surfaces are added to make all meshes water-tight, which is a requirement for volume computations. These newly introduced surfaces are shown in Figure 7.1 as colored line segments parallel to the cutting plane (outer box). Therefore, when computing distance distributions, the closest point of a cell membrane to a mitochondrion is often a boundary surface. This distorts the distance distributions and limits their interpretability. That is why we add a margin region to the cutting plane of the volume by adding an offset to it. In this margin region, distance distributions are not calculated. Figure 7.1 illustrates the margin region as the space between the outer and the inner box.

7.2.2 Input Data Specification

Neuroscientists often acquire their own data sets, which they might want to analyze using NeuroKit. For this, we specified a standardized data format and naming conventions to support loading arbitrary data sets into NeuroKit. This makes NeuroKit accessible to numerous experts and not only to our collaborators. The surface meshes must be specified in the wavefront data format. The mesh objects must follow a strict naming convention.

For instance, axons and dendrites are specified by respective names first followed by a three digit identification number:

```
Dendrite145
Axon033
```

The naming convention of mitochondria first specifies a mitochondria identifier followed by the respective neurite (axon or dendrite) and an identification number:

```
Mito_D145_01
Mito_A033_03
```

The above names show the convention for mitochondrion 1 of dendrite 145 and mitochondrion 3 of axon 33. Similarly the identifiers for boutons and spines are defined:

```
Spine_D145_02
Bouton_A033_03
```

The naming convention for synapses is composed by the synapse identifier followed by the identifier of the respective dendrite, spine, axon, and bouton:

```
Syn_D145s02A033b03
```


Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

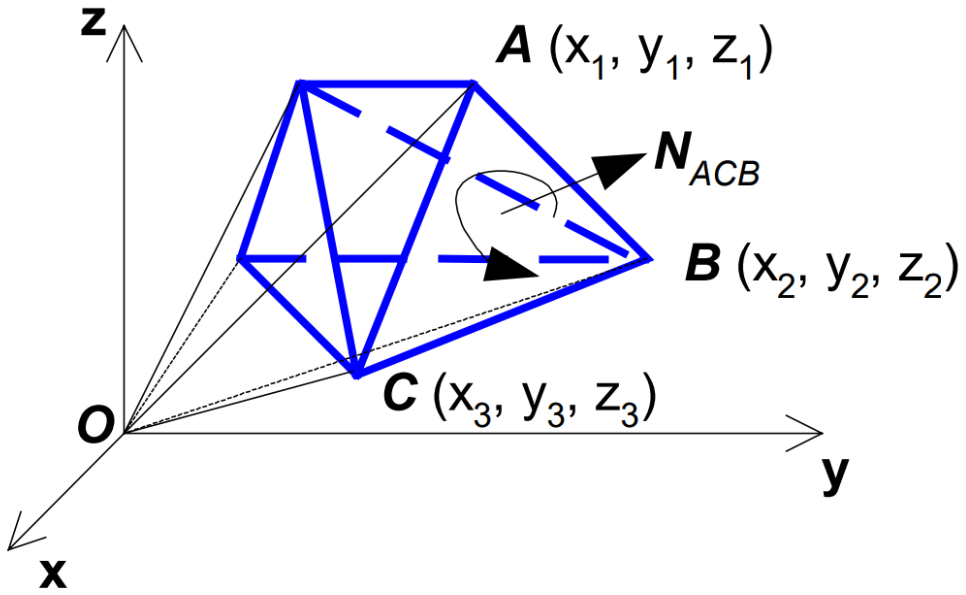


Figure 7.2: **Volume computation without transforming the meshes into a volumetric data representation.** Instead, the mesh volume can be calculated by the volume of the tetrahedron each triangle forms with the origin of the coordinate space. Source: [ZC01]

Here we see the name for a synapse between dendrite 145, spine 2, axon 33, and bouton 3. Additionally, the surface must be triangulated. We use this standardized naming convention to derive relationships between the neuronal structures, which is important to identify their neighborhoods. Additionally, EM image volumes are specified in the .raw data format with known data types and dimensions. More information about the NeuroKit data format can be found on GitHub [Tro21].

7.2.3 Volume Computation

Another mesh attribute that is of interest for the neurobiologists is the volume of mitochondria. We use an algorithm that computes the volume of the space enclosed by meshes without transforming them into a volumetric data representation [ZC01]. This decreases computation time and saves memory space. Figure 7.2 shows a mesh from which we want to calculate the volume. We assume that the mesh is triangulated. The volume computation is done in three steps. First, we compute the outward-facing normal for each triangle. The normal N of the triangle ABC is defined as the cross-product $N = (C - A) \times (B - A)$. Whether the normal is facing inside the mesh or outside the

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

mesh is determined by the triangle winding order. Second, we compute the volume of the tetrahedron that each triangle forms with the origin O. For example, the volume of the tetrahedron $OACB$ is computed like

$$|V_{OACB}| = |\frac{1}{6}(-x_3y_2z_1 + x_2y_3z_1 + x_3y_1z_2 - x_1y_3z_2 - x_2y_1z_3 + x_1y_2z_3)| \ , \tag{7.1}$$

where $A = (x_1, y_1, z_1)$, $B = (x_2, y_2, z_2)$, and $C = (x_3, y_3, z_3)$. Third, the sign of the tetrahedron $OACB$ is then computed as $sgn(\vec{OA} \cdot N_{ACB})$. Triangle normals pointing away from the origin will get a positive sign wheres triangle normals pointing in the direction of the origin will have a negative sign. The volume V_{total} of the mesh can be easily computed by the sum of tetrahedra volumes V_i like

$$V_{total} = \sum_i V_i \tag{7.2}$$

Summing up the signed volumes ensures that all parts of V_i that are not inside the mesh cancel out.

7.3 Implementation

NeuroKit is implemented in C++ 17 using Qt5 for the user interface and OpenGL 4.3 for real-time 3D rendering. The analysis views use d3.js [BOH11] and phylotree.js [SWKP18]. We have implemented a state-of-the-art real-time transparency rendering algorithm using per-pixel linked lists [Wol11] for the rendering of semi-transparent surfaces. We use Qt's WebChannel class to send data objects from C++ to the JavaScript frontend of our application and vice versa. The user study and all tests were performed on a Windows 10 machine, using an AMD Ryzen 9 3900X CPU, an NVIDIA GeForce RTX 2080 Super GPU, and 128 GB RAM. NeuroKit requires a GPU for interactive 3D rendering. NeuroKit is open-source and available on Github [Tro21].

7.3.1 Realtime 3D Rendering

NeuroKit uses detailed real-time 3D renderings to visualize neighborhoods of neuronal structures. We render the surfaces of nerve cells, cell organelles, and synapses reconstructed from dense segmentations of EM volumes of brain tissue.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

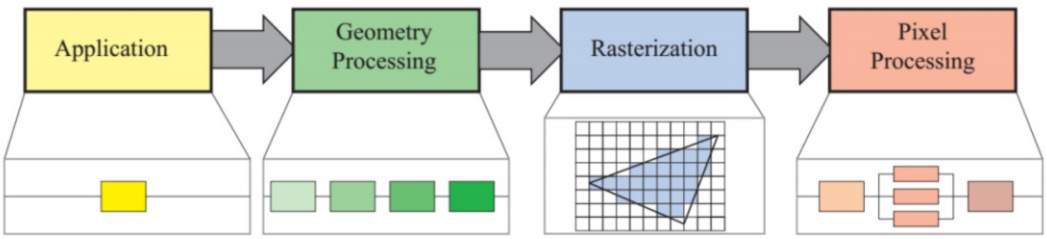


Figure 7.3: The real-time rendering pipeline. Source: [Wim20].

The Rendering Pipeline

The rendering pipeline [Wim20] describes all necessary steps to transform three-dimensional surface data to rasterized two-dimensional images on the screen. Figure 7.3 shows a simplified version of the rendering pipeline. The **application** controls issuing commands to the graphics card (GPU). Therefore, modern graphics cards offer interfaces such as OpenGL, DirectX, or Vulkan. In the **geometry processing** stage, geometry is modified in various sub-steps, and primitives are assembled. For instance, geometry primitives are points, lines, triangles, or patches. The respective primitives assemble the rendered geometry. In the **rasterization** stage, the primitives, defined in continuous space, are discretized to a finite set of pixels. This process is called sampling. Also, vertex attributes such as color values or normals are interpolated. In the final **pixel-processing** step, per-fragment operations are performed. Fragments are pixel candidates. For instance, the color of a pixel is determined by multiple fragments of different objects if the scene contains transparent objects. This process is called blending. Other per-fragment operations include fragment shading, stencil testing, or dithering.

Order Independent Transparency

Correct transparency renderings are essential to observe how cell organelles are positioned within neurons. Since simple blending techniques require complex triangle depth sorting to display transparency correctly, we implemented a state-of-the-art order-independent transparency algorithm. Rendering objects translucently means that multiple objects may compose the color of each pixel. This approach stores all color, alpha, and depth values of each fragment candidates in a linked list for each pixel. In the next step, each pixel list is sorted by the fragments' depth such that the fragment closest to the viewing plane is the first element, and the furthest fragment is the last element in the list. After this step, all fragments are blended, which results in the final pixel value. Figure 7.4 shows the results of this algorithm with our data.

7.3.2 Analysis Definition Files

In the following section, we explain the functionality of the analysis definition files (ADFs) in more detail. ADFs are used to customize visualizations and to allow the user to specify

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

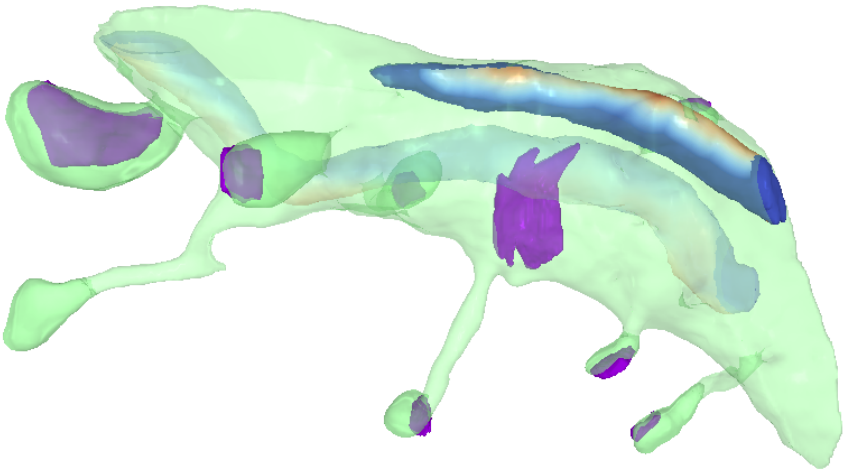


Figure 7.4: 3D rendering using per-pixel linked lists for order independent transparency (OIT).

domain-specific usage scenarios. ADFs are specified in the human-readable JSON data format. First, we give an overview of the general structure and the implemented design space and then give two concrete examples of ADFs.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Design Space Overview

The JSON snippet in Listing 7.1 illustratesillustrates the general structure of an ADF. Each ADF specifies a usage scenario. Each usage scenario can consist of multiple subtasks/domain tasks, whereby each domain task specifies up to three visualization methods at different cardinalities (*low-cardinality visualizations*, *medium-cardinality visualizations*, *high-cardinality visualizations*). Each visualization method is best suited to visualize and compare a certain number of spatial neighborhoods.

```
{
    "name": "usage scenario name",
    "subtasks": [
        {
            "name": "domain task name",
            "low_cardinality_vis": { ... },
            "medium_cardinality_vis": { ... },
            "high_cardinality_vis": { ... }
        },
        { ... }
    ]
}
```

Listing 7.1: General structure of an ADF file

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

In Listing 7.2, we explain how the individual visualization methods are specified. First, the **scatter-plot visualization**, is specified by the *scatter-plot* keyword. Based on the selected target structure, data points in the scatter-plot represent either *synapses* or *mitochondria*. We plan to allow more target structures in future versions of NeuroKit. The specification also allows us to map different data attributes to the axes of the scatter-plot. Detailed explanations of the data attributes are available in the NeuroKit wiki on Github [Tro21].

```
{
    "id": 0,
    "name": "scatter-plot",
    "target-structure": ["mitochondria", "synapses"],
    "x-axis": {
      "label": "here goes the x-axis label",
      "attribute": [ "mito-min-distance-to-cell",
                     // minimal distance of mitochondrion
                     // to the cell membrane
                     "mito-volume",
                     // mitochondrion volume
                     "mito-spine-coverage",
                     // number of spines covered
                     // by the mitochondrion
                     "mito-surface-closer-than-threshold",
                     // percentage of the mitochondrion surface
                     // closer to the cell membrane
                     // than a user defined threshold
                     "syn-surf-area",
                     // surface area of a synapse
                     "syn-close-mitos"
                     // number of synapses that are considered
                     // close to a mitochondrion
                     // by a user-defined threshold ]
    },
    "y-axis": {
      "label": "here goes the y-axis label",
      "attribute": [ "mito-min-distance-to-cell", "mito-volume",
                     "mito-spine-coverage",
                     "mito-surface-closer-than-threshold",
                     "syn-surf-area", "syn-close-mitos"]
    }
}
```

Listing 7.2: ADF structure of a scatterplot visualization

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

The **distance-tree visualization** is defined by the *distance-tree* keyword (see Listing 7.3). It supports both mitochondria and synapses as target structures. A couple of the parameters are used to fine-tune data mappings.

```
{
    "id": 1,
    "name": "distance-tree",
    "target-structure": ["mitochondria", "synapses"],
    "params": ["related-synapses", "surrounding-synapses",
    "surrounding-mitochondria"]
}
```

Listing 7.3: ADF structure of a distance-tree visualization

The ADF defines **grouped bar charts** as illustrated below (see Listing 7.4).

```
{
    "id": 2,
    "name": "grouped-bar-chart",
    "target-structure": ["mitochondria", "synapses"],
    "params": ["surrounding-synapses", "related-synapses",
    "surrounding-mitochondria"]
}
```

Listing 7.4: ADF structure of a grouped bar chart visualization

The **distance-matrix** visualization is defined like (see Listing 7.5)

```
{
    "id": 3,
    "name": "distance-matrix",
    "target-structure": ["mitochondria"],
    "params": ["related-synapses", "surrounding-synapses"]
}
```

Listing 7.5: ADF structure of a distance-matrix visualization

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

The **cummulative-histogram** (see Listing 7.6) allows mapping data attributes to the bins and makes the number of bins adjustable by the *adjustable* parameter.

```
{
    "id": 4,
    "name": "cummulative-histogram",
    "target-structure": ["mitochondria"],
    "bins": "distance-to-cell-membrane",
    "params": ["fixed", "adjustable"]
}
```

Listing 7.6: ADF structure of a cummulative-histogram visualization

Violin-plots can be normalized through the respective parameter in the ADF (see Listing 7.7). By normalizing a violin plot, its shape gets mapped to an area of one, allowing comparison of distance distributions of cell organelles with significant size differences.

```
{
    "id": 5,
    "target-structure": "mitochondria",
    "name": "violin-plot",
    "bins": "distance-to-cell-membrane",
    "normalized": true
}
```

Listing 7.7: ADF structure of a violin-plot visualization

The **dendrite abstraction level 1** (see Fig. 8.6a, and Listing 7.8) creates a visualization inspired by Neurolines [AABS+14]. Additionally, we allow to encode synapse properties onto spine abstractions using the respective parameter value.

```
{
    "id": 6,
    "target-structure": "mitochondria",
    "name": "dendrite-abstraction-1",
    "param": "encode-spine-sa"
}
```

Listing 7.8: ADF structure of the dendrite abstraction level 1 visualization

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

We specify the **dendrite abstraction level 2** (see Fig. 8.6b, and Listing 7.9) as illustrated below. It currently does not allow custom user specifications. However, we plan to allow normalization of the dendrite abstractions in a future release of NeuroKit.

```
{
    "id": 7,
    "target-structure": "mitochondria",
    "name": "dendrite-abstraction-2"
}
```

Listing 7.9: ADF structure of the dendrite abstraction level 12 visualization

Examples

We present two concrete examples of analysis definition files used and adjusted by domain experts during the evaluation of NeuroKit (Section 9.1). The listing 7.10, 7.11, 7.12 show the ADFs of the usage scenario *mitochondria analysis* (Chapter 8). For this scenario, we specified three subtasks/domain tasks. They illustrate how each of the domain tasks defines three visualizations with varying cardinalities. The study participants fine-tuned the data visualizations by adjusting parameters and data mappings. Listing 7.13 shows another usage scenario comparing two data sets of mice at a different age. This scenario only specifies one domain task but defines synapses as a target structure, in contrast to the ADF shown in the Listings 7.10, 7.11, 7.12. Figure 1.1 shows NeuroKit visualizing data of this scenario.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

```
[
  {
    "name": "Mitochondria Analysis",
        "datasets": ["mouse3"],
    "subtasks": [
      {
        "name": "DT 1: synaptic neighborhood of mitochondria",
        "low_cardinality_vis": {
          "id": 0,
          "target-structure": "mitochondria",
          "name": "distance-tree",
          "params": "related-synapses"
        },
        "medium_cardinality_vis": {
          "id": 1,
          "target-structure": "mitochondria",
          "name": "grouped-bar-chart",
          "params": "related-synapses"
        },
        "high_cardinality_vis": {
          "id": 2,
          "target-structure": "mitochondria",
          "name": "distance-matrix",
          "params": "related-synapses"
        }
      },
      [...]
      ]
    }
]
```

Listing 7.10: **An example of an analysis definition file (ADF).** The usage scenario *Mitochondria Analysis* uses mitochondria as the target structure. We describe the usage scenario *Mitochondria Analysis* in detail in Chapter 8 of the thesis. It specifies one domain tasks (DT1: Synaptic neighborhood of mitochondria). he tasks supports visual comparison at three cardinality levels by defining a specific visualization method for each level. The visualization methods define a set of parameters and data mappings.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

```
[ { [...],
    "subtasks": [
      {
        "name": "DT 2: mitochondria placement",
        "low_cardinality_vis": {
          "id": 3,
          "target-structure": "mitochondria",
          "name": "cumulative-histogram",
          "bins": "distance-to-cell-membrane",
          "normalized": false,
          "params": "adjustable"
        },
        "medium_cardinality_vis": {
          "id": 4,
          "target-structure": "mitochondria",
          "name": "violin-plot",
          "bins": "distance-to-cell-membrane",
          "normalized": true
        },
        "high_cardinality_vis": {
          "id": 5,
          "target-structure": "mitochondria",
          "name": "scatter-plot",
          "x-axis": {
            "label": "Minimal Distance to Cell Boundary",
            "attribute": "min-distance-to-cell"
          },
          "y-axis": {
            "label": "mitochondria volume in cubic microns",
            "attribute": "mito-volume"
} } } ] } ]
```

Listing 7.11: **An example of an analysis definition file (ADF).** It uses mitochondria as the target structure. It specifies one domain task (DT2: Mitochondria placement). The subtask supports visual comparison at three cardinality levels by defining a specific visualization method for each level. The visualization methods define a set of parameters and data mappings.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

```
[
  {
    [...],
    "subtasks": [
      {
        "name": "DT 3: spine coverage of mitochondria",
        "low_cardinality_vis": {
          "id": 6,
          "target-structure": "mitochondria",
          "name": "dendrite-abstraction-1",
          "param": "encode-spine-sa"
        },
        "medium_cardinality_vis": {
          "id": 7,
          "target-structure": "mitochondria",
          "name": "dendrite-abstraction-2"
        },
        "high_cardinality_vis": {
          "id": 8,
          "target-structure": "mitochondria",
          "name": "scatter-plot",
          "x-axis": {
            "label": "number of spines covered by mitochondrion",
            "attribute": "mito-spine-coverage"
          },
          "y-axis": {
            "label": "mitochondria volume in cubic microns",
            "attribute": "mito-volume"
          }
        }
      }
    ]
  }
]
```

Listing 7.12: **An example of an analysis definition file (ADF).** It uses mitochondria as the target structure. It specifies one domain task (DT3: Spine coverage of mitochondria). The subtask supports visual comparison at three cardinality levels by defining a specific visualization method for each level. The visualization methods define a set of parameters and data mappings.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

```
{
  "name": "Relation of synapses and mitochondria",
  "datasets": ["mouse2", "mouse3"],
  "subtasks": [
    {
      "name": "DT 1: surrounding mitochondria",
      "low_scale_vis":
      {
        "id": 9,
        "target-structure": "synapses",
        "name": "distance-tree",
        "params": "surrounding-mitochondria"
      },
      "medium_scale_vis": {
        "id": 10,
        "target-structure": "synapses",
        "name": "grouped-bar-chart",
        "params": "surrounding-mitochondria"
      },
      "high_scale_vis": {
        "id": 11,
        "target-structure": "synapses",
        "name": "scatter-plot",
        "y-axis": {
          "label": "Synapse surface area",
          "attribute": "syn-surf-area"
        },
        "x-axis": {
          "label": "closest mito surface area",
          "attribute": "syn-close-mitos"
        }
          }
    }
  ]
}
```

Listing 7.13: **An example of an analysis definition file (ADF) with multiple data sources.** The usage scenario *Relation of synapses and mitochondria* defines one subtask (DT1: Surrounding mitochondria) and specifies synapses as target structures. In this task, scientists can compare the synapses of two different data sets, to investigate neuronal development over time.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

CHAPTER 8

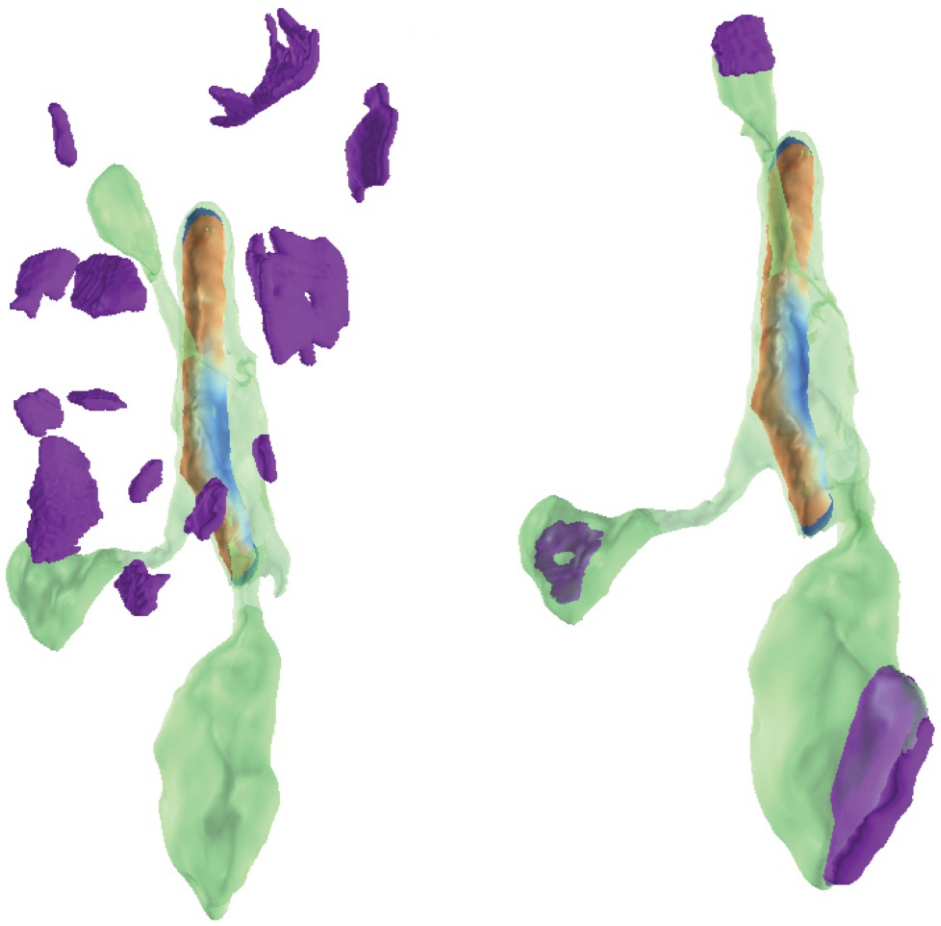
Use Case: Mitochondria Analysis

One of our collaborator's main research focus are mitochondria, their spatial neighborhood, and their relation to synapses. We describe the three domain tasks that comprise the user-defined analysis scenario *Mitochondria Analysis* (see Fig. 8.2, 8.4, and 8.6) and the novel visual encodings we designed for it. The corresponding analysis definition file is shown in Listing 7.10. The user study in Section 9.1 focuses on the analysis scenario presented here. Prior to NeuroKit, our collaborators had no means to analyze mitochondria neighborhoods other than looking at the raw 3D data and running Python scripts.

8.1 Synaptic Neighborhood of Mitochondria

8.1.1 Domain Goal

The goal of the first domain task in the mitochondria analysis scenario is to explore the spatial relationships between mitochondria and synapses. The synaptic neighborhood of mitochondria might give a detailed insight into the role mitochondria play during synaptogenesis or synapse potentiation. There are two aspects to explore. First, our collaborators want to look at the synaptic densities that correlate spatially with the selected mitochondrion's cell. Second, they also want to analyze all synapses within a certain distance to the mitochondrion, regardless of whether they are connected to the current neurite or not. This second aspect is part of an initial hypothesis of our collaborator and has not been evaluated in-depth before. Figure 8.1 illustrates these two different aspects.



(a) Surrounding synapses (b) Related synapses

Figure 8.1: **Two different synaptic neighborhoods of mitochondria.** (a) all synapses that are within 1 micron from the mitochondrion and (b) synapses connected to the mitochondrion's dendrite.

8.1.2 Realization

To support the analysis of this domain task, in collaboration with our domain experts, we designed the three comparative visualizations shown in Fig 8.2. These views allow the scientist to explore mitochondria and synapse relationships at different scales (i.e., comparing few to many mitochondria) and intuitively navigating between them.

In the **low-cardinality** comparative visualization we use small multiples to show a synaptic distance tree (Fig 8.2a) for each mitochondrion under investigation. The distance tree is an adaption of a phylogenetic tree (Fig 8.2a). For each mitochondrion under investigation, we create a synaptic distance tree. The root of the tree represents

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

the mitochondrion, and the root's children represent synapses. The horizontal edge length from the root to a child indicates the distance between the mitochondrion and the synapse. Additionally, in a third layer of the tree, we encode the distance from the synapse (Layer 2) to any mitochondria in the connecting cell (Layer 3). This view allows an in-depth comparison of the number of synapses and their distances to each selected mitochondrion. The **medium-cardinality** view shows a sorted bar chart for each mitochondrion, where each bar represents the distance to an individual synapse (Fig 8.2b). This view is more compact but still shows distances to synapses and the overall distance distribution. The **high-cardinality** view gives scientists a general overview and supports spotting outliers and see trends quickly. We show a distance matrix to compare dozens to hundreds of mitochondria (Fig 8.2c). Each row in the matrix corresponds to one particular mitochondrion, and each column represents a specific synapse. Distances between mitochondria and synapses are color-coded, darker colors representing closer distances.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

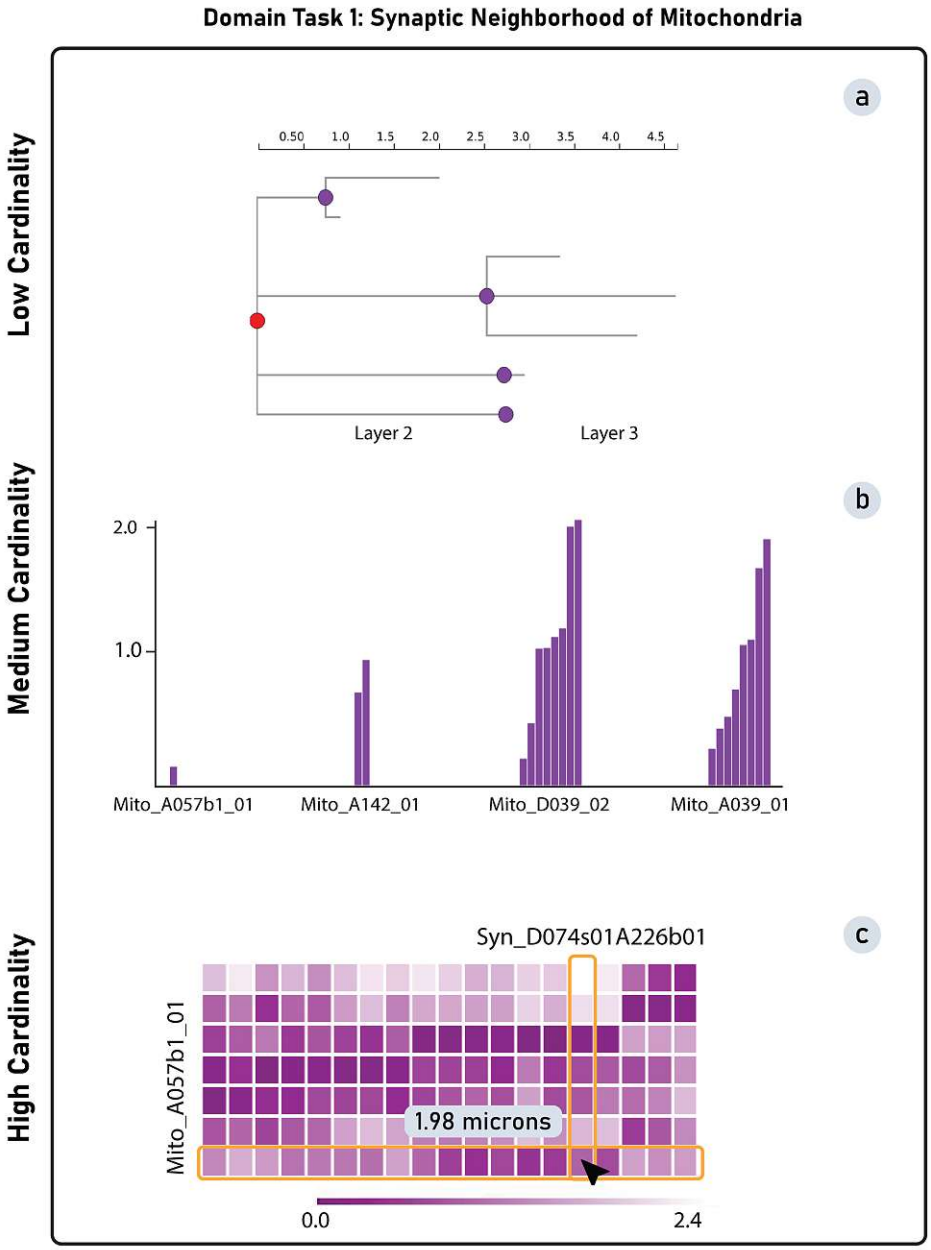
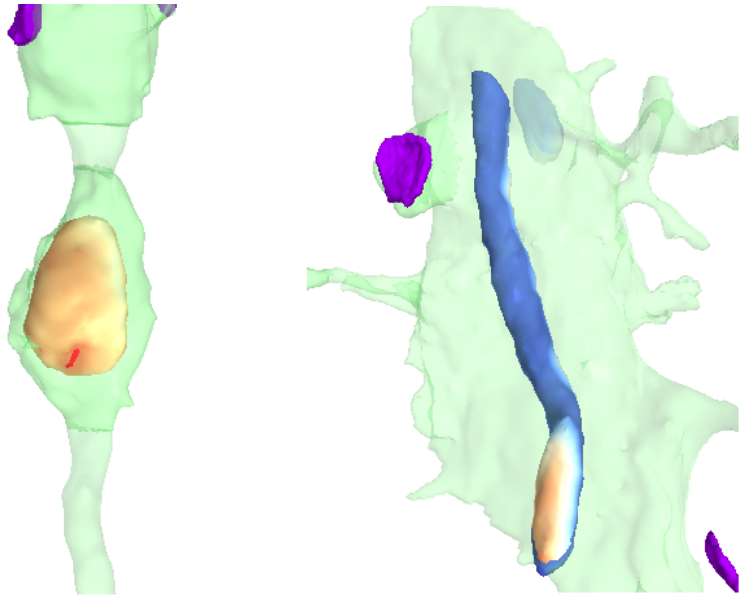


Figure 8.2: **Analysis views for the domain task *Synaptic Neighborhood of Mitochondria*.** (a) - (c): low-, medium-, high-cardinality comparison. (a) distance tree, (b) grouped bar chart, (c) distance matrix.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) Axonal mitochondrion (b) Dendritic mitochondrion

Figure 8.3: **Two mitochondria placed differently in their cells.** The axonal mitochondrion (a) is much closer to the cell membrane, than the dendritic mitochondrion (b). A red color on the mitochondrions surface indicates close proximity while blue represents a greater distance.

8.2 Mitochondria Placement

8.2.1 Domain Goal

Our collaborators want to investigate the spatial relationships between mitochondria and the cell membrane in the second domain task. The scientists want to investigate how close mitochondria come to the cell membrane and if cell regions with proximity to the mitochondrion differ from other areas. Since mitochondria provide energy to the cell, such morphological influences are of high interest. For this task, scientists need to look at the detailed 3D position of a mitochondrion and want to see the percentage of a mitochondrion's surface close to the cell boundary.

8.2.2 Realization

To show mitochondria locations within a cell, we color-code each mitochondrion in the 3D view to encode the distance from each surface point of the mitochondrion to the closest point of the cell membrane. Red regions indicate close proximity, while blue areas reveal regions more distant to the cell membrane (Fig. 6.2, 8.3). In addition to the color-coding in 3D, we have designed three analysis views that support a comparative analysis at different

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

cardinality levels (Fig 8.4). In the **low-cardinality** view, we show small multiples of cumulative histograms for each mitochondrion. We plot the distances from each surface point of the mitochondrion to the closest point of the cell membrane (Fig 8.4a). This gives the overall distribution of distances, indicating whether a mitochondrion follows closely along the cell membrane or only comes close in a small area. In the **medium-cardinality** visualization, we depict a series of violin plots next to each other (Fig 8.4b). Each shape corresponds to a single mitochondrion and allows the comparison between different mitochondria. The vertical axis represents the distance between a mitochondria surface point and the cell membrane. The width of the violin shape represents how much of the mitochondrion's surface has a certain distance to the cell membrane. The thicker the shape is at a specific height, the more of the mitochondrion's surface is within that distance to the cell membrane. With this view, scientists can quickly judge the overall distance distribution for a single mitochondrion and compare the distributions of different mitochondria. The **high-cardinality** view shows a scatterplot where each dot represents a single mitochondrion (Fig 8.4b). The horizontal axis represents the distance from the mitochondrion's closest point to the cell membrane. On the vertical axis, we encode the mitochondrion's volume.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

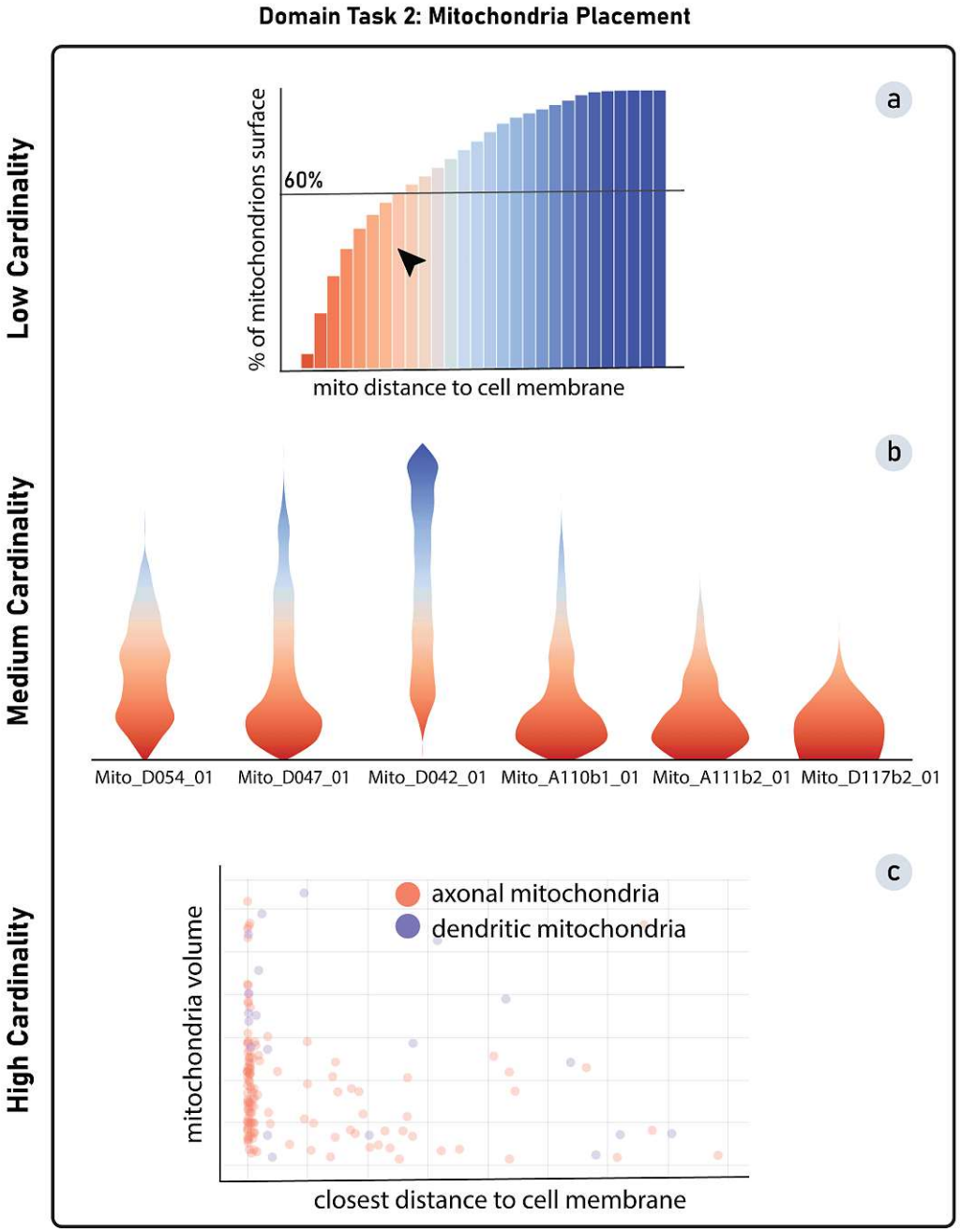


Figure 8.4: **Analysis views for the domain task *Mitochondria Placement*.** (a)-(c): low-, medium-, high-cardinality comparison. (a) cumulative histogram, (b) violin plot, (c) scatter plot.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

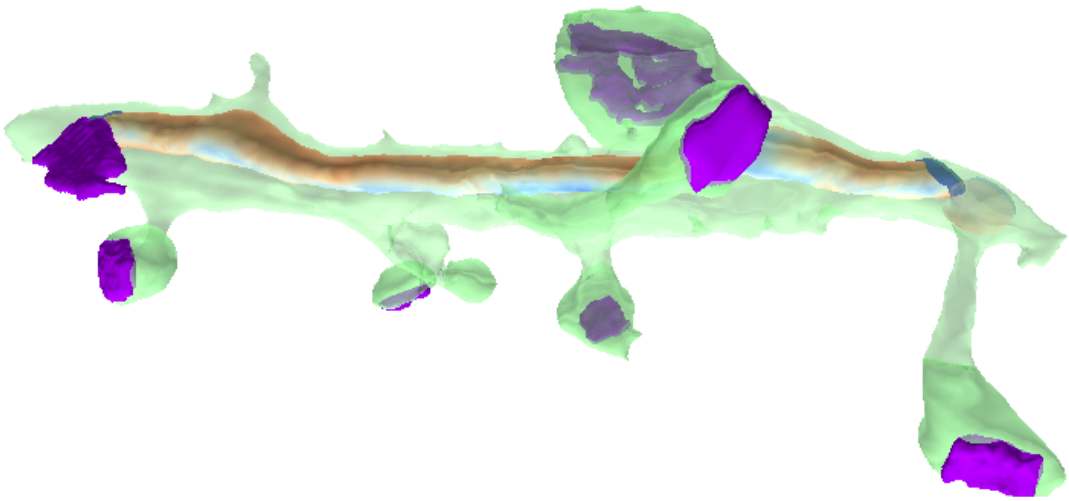


Figure 8.5: **Spine coverage.** All spines (tiny branches) are covered by a mitochondrion.

8.3 Mitochondric Spine Coverage

8.3.1 Domain Goal

In the third part of the mitochondria analysis, our collaborators study the relationship between dendritic spines and mitochondria. Spine coverage is defined as a mitochondrion being present in the cell region right below the neck of a spine or dendritic shaft (see Fig. 8.5). This property is of high biological interest because mitochondria might influence spine plasticity, which may influence the fission/fusion dynamics of mitochondria.

8.3.2 Realization

The three comparative visualizations for this task can be seen in Fig. 8.6. In the **low-cardinality** view (Fig 8.6a), we designed a novel view to highlight dendritic spines in each selected neighborhood. First, we abstract the dendrite to a straight horizontal line in 2D, inspired by Neurolines [AABS+14]. Next, we encode the mitochondrion position (purple shapes) within the dendrite and represent spines and their lengths as green vertical lines. The disks at the top of the spine represents a synapse, and its radius encodes the spine's surface area to volume ratio, which gives an intuition of the spine's shape. This view provides a detailed analysis of the morphology of the dendrites (i.e., mitochondria positions, spine locations, and synapse sizes). The **medium-cardinality** comparative visualization (Fig. 8.6b) represents each dendrite as a vertical rectangle, to be able to fit more dendrites onto the screen for comparisons (Fig 8.6b). The rectangle's height represents the length of the dendrite, and the positions of the mitochondria within the dendrite are visualized through smaller pink rectangles. Spine positions are indicated

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

as black horizontal lines. The **high-cardinality** comparative visualization shows a scatterplot (Fig 8.6c). Each data point corresponds to an individual mitochondrion. The horizontal axis encodes the number of spines covered by this mitochondrion, and the vertical axis encodes the mitochondrion's volume. Figure 8.7 and Figure 8.8 show two screenshots of the user interface of NeuroKit in the usage scenario *mitochondria analysis*.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

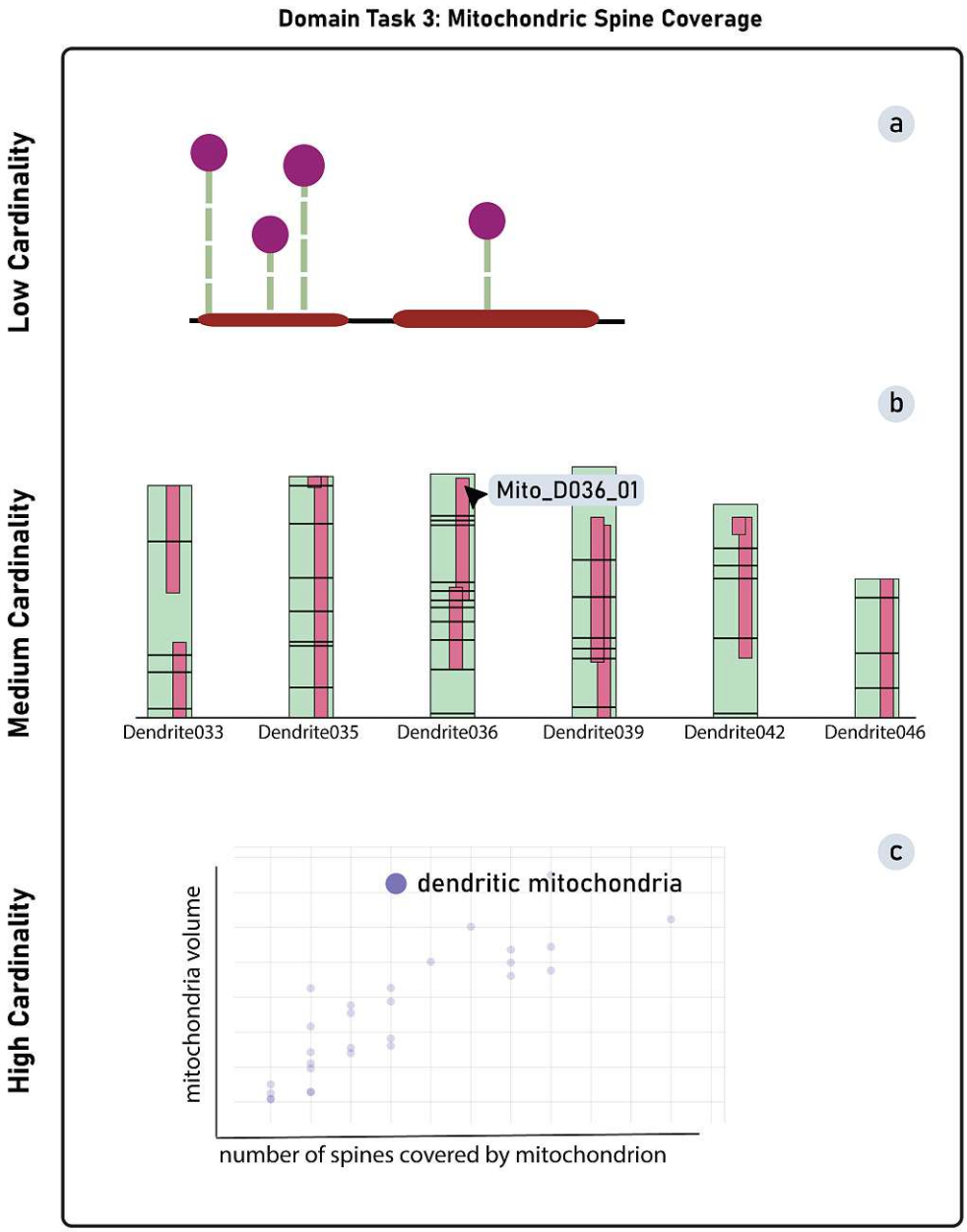


Figure 8.6: **Analysis views for the domain task *Mitochondric Spine Coverage*.** (a) - (c): low-, medium-, high-cardinality comparison. (a) dendrite abstraction level 1, (b) dendrite abstraction level 2, (c) scatter plot.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

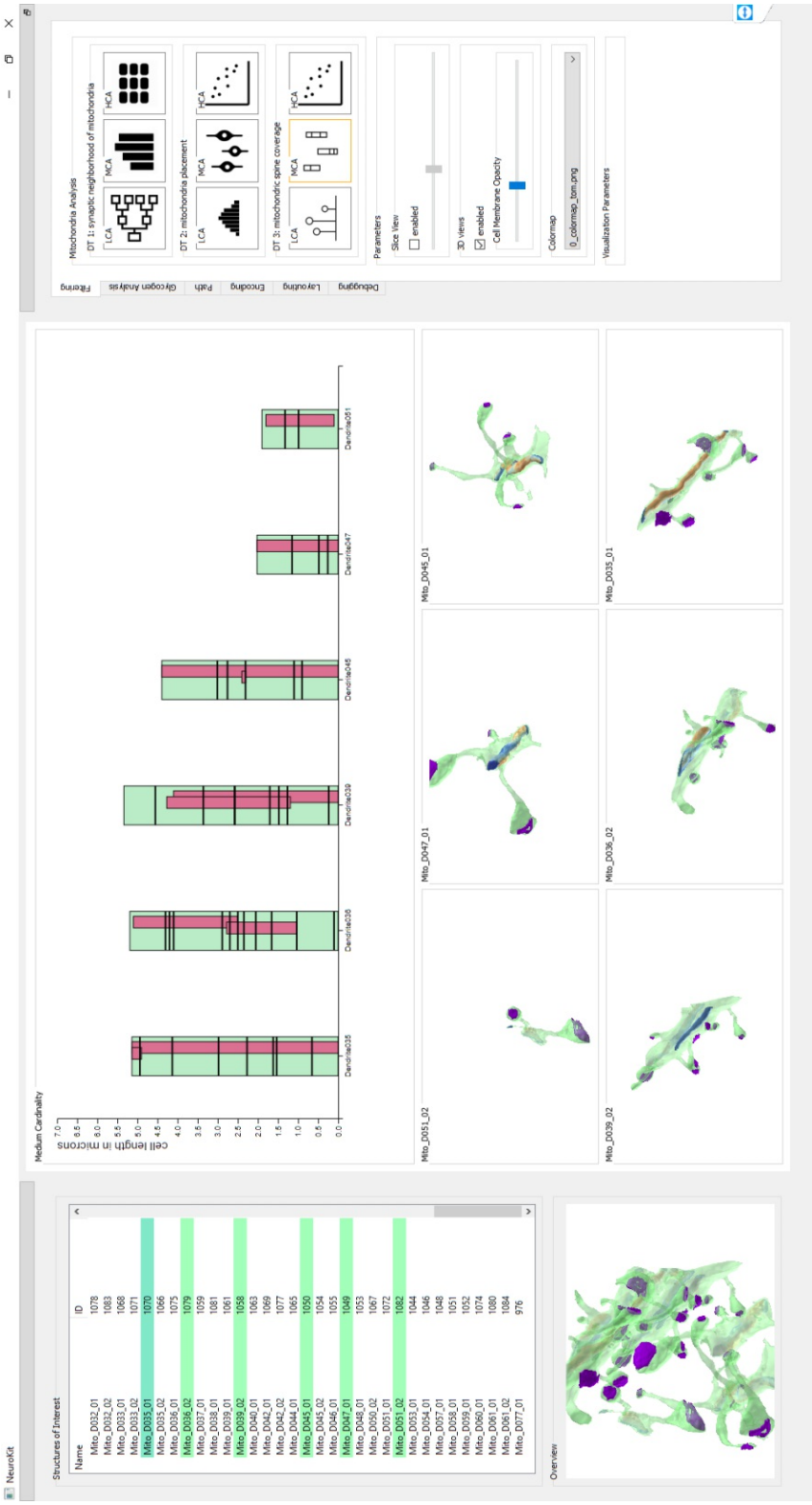


Figure 8.7: **Screenshot of NeuroKit s' user interface** showing the domain task *Mitochondric Spine Coverage* in the medium cardinality visualization (dendrite abstraction level 2).

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

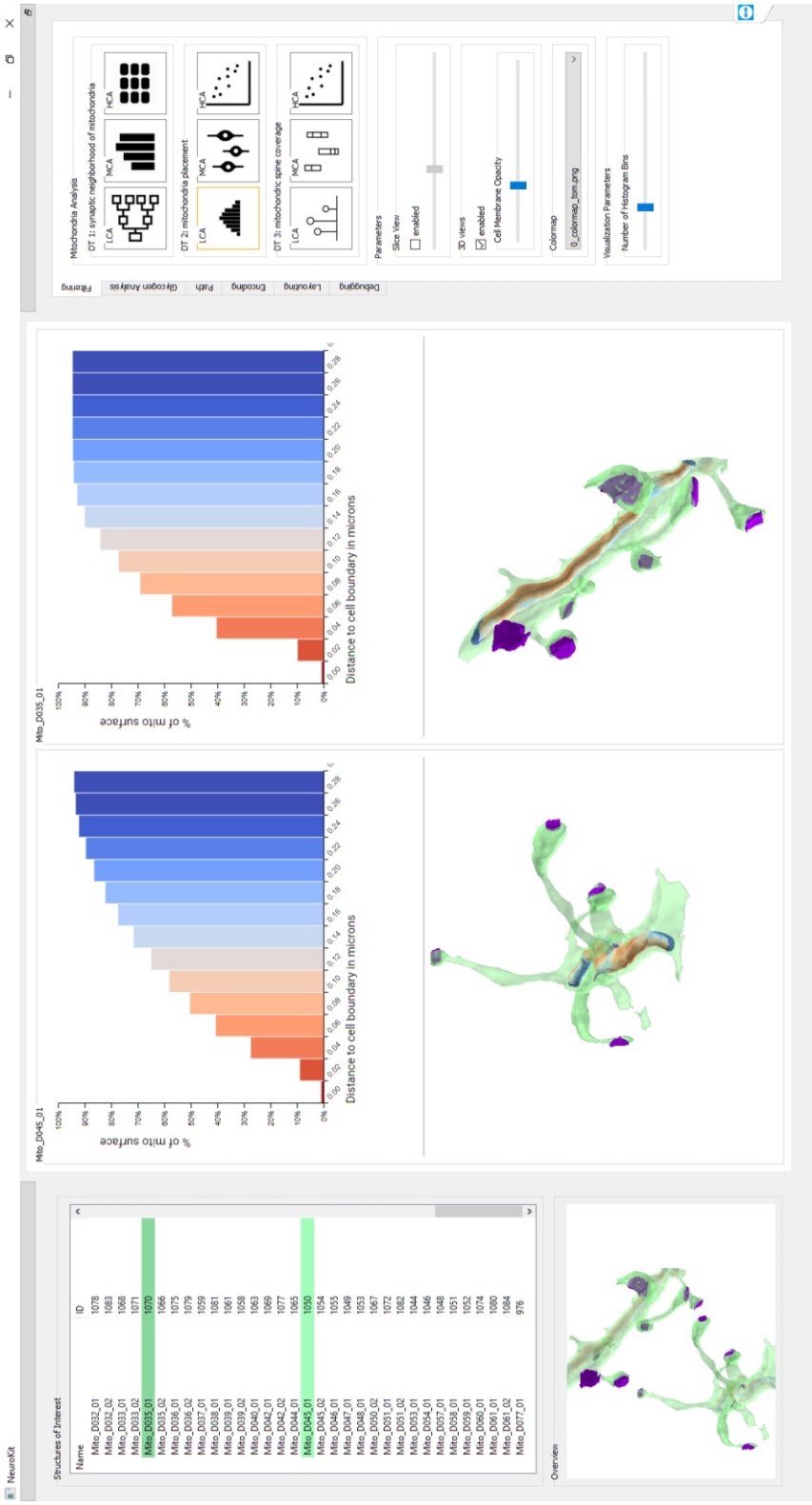


Figure 8.8: **Screenshot of NeuroKit s' user interface** showing the domain task *Mitochondria Placement* in the low cardinality visualization (cumulative histogram).

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

CHAPTER 9

Evaluation

We report on a qualitative user study of the mitochondria analysis scenario with eight domain experts to assess the effectiveness and usefulness of NeuroKit and present key findings of our evaluation. We held a exploratory feedback sessions with the study participants and share the results of a survey. We conducted exploratory sessions based on our mitochondria analysis scenario, collected think-aloud feedback, and asked study participants to fill out feedback surveys after each completed domain task.

9.1 User Study: Mitochondria Analysis

Participants

We evaluated our application with eight domain experts, of which one was male, and seven were female. The participants are affiliated with the Neuroscience Institute at the University of Turin and the Center for Brain Science at Harvard University. All experts are experienced in analyzing high-resolution EM data of mice brains and included three professors, one postdoctoral researcher, two senior Ph.D. students, and two undergraduate students.

Setup

We met with each participant for a two-hour session remotely over the Zoom video conferencing platform. After a 30 min introduction to NeuroKit and its user interface, we started a data exploration session. Due to our remote setup, we followed a collaborative visual analytics evaluation protocol by Arias-Hernandez et al. [AHKGF11], where the domain experts guided and steered the analysis while we operated the user interface. This allowed participants to focus on domain-specific questions rather than issues related

to the user interface. We collected think-aloud feedback and written survey responses after each completed domain task.

In the user study, we utilized one of the six data sets provided by the domain experts showing layer one of the somatosensory cortex of a four-month-old mouse with 1088 segmented structures (167 synapses, 152 spines, 237 boutons, 81 dendrites, 274 axons, and 177 mitochondria). We asked participants to perform the three domain tasks of the mitochondria analysis scenario described in Chapter 8. Prior to NeuroKit, our collaborators used Blender for their visual exploration and had to manually mark and compute distances between structures of interest. They did not have any support for visual comparisons other than taking screenshots.

Analyzing Synaptic Neighborhoods of Mitochondria

In this task, participants explored the spatial relationships between mitochondria and surrounding synapses (see Sec. 8.1 and Fig. 8.2). Analyzing the synaptic neighborhood of mitochondria is crucial for understanding the mitochondria's role in support of the high energy-consuming brain regions like synapses. Therefore, it is interesting to investigate how synapses are distributed around mitochondria. Participant P2 configured NeuroKit such that it takes synapses within a certain threshold to the mitochondrions' surface into account (see Fig. 8.1b). P2 started the session by analyzing a single dendritic mitochondrion in the synaptic distance tree view (low-cardinality visualization). He discovered two particularly close synapses to this mitochondrion. He selected these two synapses in the distance tree view and examined them in more detail in the neighborhood view. While analyzing the 3D view, he discovered the proximity between the selected mitochondrion and the cell membrane. The combination of analysis views and neighborhood view helped P2 to form the hypothesis that the two synapses and the mitochondrion's region close to the cell membrane might interact. P2 started to select more dendritic and axonal neighborhoods of mitochondria.

P2 spotted a bouton containing a small mitochondrion with a nearby synapse while analyzing the distances between multiple mitochondria and their related synapses. Our collaborators are interested in whether boutons contain synapses or not. The presence of a mitochondrion means that the bouton is mature and has undergone some potentiation. P2 could quickly tell from the bar chart analysis view that the bouton and synapse were related. Next, P2 used the high-cardinality distance-matrix to check his hypothesis that mitochondria within boutons tend to have synapses close by and spotted a boutonal mitochondrion with more distant synapses. He enabled the EM slice view for additional exploration and noted that this observation warrants further analysis.

P2 configured this domain task to take into account only synapses related to the mitochondrion's cell (see Fig. 8.1b). In contrast, P6 and P7 chose to analyze the distribution of all surrounding synapses of the mitochondria (see Fig. 8.1a). P7 further adjusted the distance threshold to fine-tune which synapses were still considered part of a neighborhood. P6 and P7 analyzed the synapse distribution around mitochondria in the

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

low- and medium-cardinality views. They stated that these visualizations could help them understand why mitochondria shapes differ in distinct parts of the nerve cell. For instance, mitochondria in the nucleus form complex nets, while dendritic mitochondria are mostly tubular structures. P3 was interested in analyzing both the related synapses to the mitochondrion and the surrounding synapses.

Analyzing Mitochondria Placement

In the mitochondria placement task, participants explored the proximity between mitochondria and their surrounding neurite membranes (see Sec. 8.2 and Fig. 8.4). First, P2 changed the dendrites' opacities to understand how the mitochondria are positioned in their cell. Next, he compared the distance distribution histograms of three dendritic mitochondria. He discovered that two mitochondria both have 60% of their surface closer than 100nm to the cell membrane, while the third mitochondrion lies much more intracellular. P6 and P7 used the histogram (low-cardinality view) and the violin plot (medium-cardinality view) to explore the different mitochondria locations in axons and dendrites. P7 selected four axonal mitochondria and three dendritic mitochondria in the medium-cardinality visualization. She discovered that axonal mitochondria tend to get very close to the cell membrane, while dendritic mitochondria tend to have regions with more space between the mitochondrion and the cell membrane. 3D neighborhood views of her analysis are shown in Fig. 1.2c. When switching to the high-cardinality scatterplot visualization, P2, P6, and P7 chose to map the mitochondrion volume onto the vertical axis and its closest distance to the cell membrane onto the horizontal axis. Using the high-cardinality visualization, P6 could confirm her hypothesis about the different mitochondria locations in dendrites and axons. In addition, P2 and P6 used the scatterplot to study outliers in more detail. P6 and P7 requested more mitochondria attributes, such as surface area, to map onto the scatterplot's different axes. P2 also asked to view the violin plots both normalized and scaled to the mitochondrion's size, since mitochondria size is another important aspect in his analysis. We have incorporated this feedback and added a parameter to the visualization, that can now be specified via the analysis definition file.

Analyzing Mitochondria Spine Coverage

In this task, participants explored the location and size of spines in relation to mitochondria (see Sec. 8.3 and Fig. 8.6). Spines often develop synapses, which are big energy consumers. Therefore, knowing more about mitochondria close to spines plays a vital role in understanding metabolic processes. While studying the dendrites' morphology in the low-cardinality view, P1 observed that spines usually develop directly next to mitochondria. She advised us to select individual spines in the analysis view. These spines were then highlighted in the 3D neighborhood view, where she could examine them more closely. P2 selected six mitochondria to compare them in the medium-cardinality visualization. In a particular case, he observed two mitochondria that overlap in the dendrite. He speculated that these two mitochondria might have been one mitochondrion

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

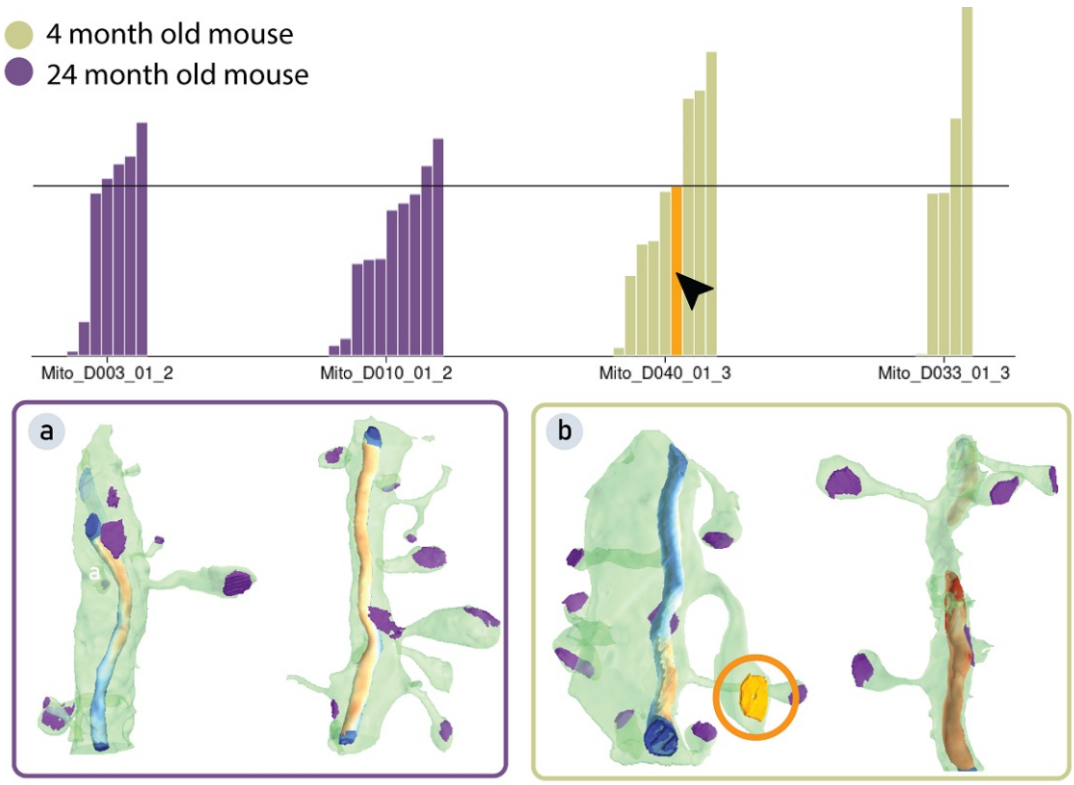


Figure 9.1: **Comparing synapse locations in two different mouse data sets.** Grouped bars show the distances between a mitochondrion and its neighboring synapses. Hovering over a bar highlights the corresponding synapse in the 3D neighborhood view. The young mouse (purple, a) exhibits *more* and *closer* synapses than the adult mouse (green, b), indicating synaptic pruning over time.

earlier that divided into two mitochondria. P2 pointed out that he wants to use NeuroKit in the future to analyze whether regions of overlapping mitochondria develop more spines. When switching to the scatterplot visualization, P1, P2, P6, and P7 immediately spotted a linear dependence between the mitochondria volume and the number of spines covered, which was a new finding for all of them.

Analyzing Synaptic Development over Time

In addition to the three domain tasks described above, one of our collaborators also wanted to compare the synaptic neighborhood of mitochondria in multiple data sets simultaneously (see Fig. 9.1). Therefore, after completing the user study and post-study questionnaire, he continued his exploratory session. He is interested in how brain structure evolves throughout the aging process and wants to analyze spine dynamics and synaptic plasticity, a process that allows the dynamic formation, elimination, or modification of spines based on learning and memory [VCDR+20]. To support this analysis scenario,

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

he first specified the use of multiple data sources in the analysis definition file. When loading multiple data sets simultaneously, NeuroKit automatically adjusts the views to highlight which data set a data point comes from, using distinct colors for each data set. Fig. 9.1 shows a comparison between the synaptic neighborhoods of a young adult (4 months old) and a geriatric (24 months old) mouse. Our collaborator analyzed and compared four mitochondria neighborhoods with the medium-cardinality visualization. The different data sets are depicted in purple and green, respectively. Our collaborator identified that the younger mouse (purple) exhibits more and closer synapses than the older mouse (green), suggesting synaptic pruning over time. The bar chart helped the collaborator to decide whether a mitochondrion is responsible for providing energy to a specific synapse or not, which is important for synapse development.

9.2 Qualitative Feedback

Fig. 9.2 shows user ratings for general real-world applicability, tool effectiveness, and relevance for each of the three domain tasks. Seven out of eight participants rate questions Q1-Q4 (effectiveness) positively, with the remaining participant being neutral. Interestingly, while participants agreed on the effectiveness of the tool, their opinions diverged more concerning Q5, asking about the relevance of mitochondria analysis for their personal research (Fig. 9.2b). This further demonstrates the importance of our customization approach, where neuroscientists can define their own analysis scenarios within NeuroKit, either focusing on mitochondria or other target structures.

Finally, all participants agreed that our tool improved the analysis of high-resolution EM data. This is not surprising; other than NeuroKit, none of the study participants currently have visual tools available to analyze spatial neighborhoods of brain tissue. Prior to NeuroKit, P2 used Blender for his visual exploration and had to mark and compute distances between structures of interest manually. He did not have any support for visual comparisons other than taking screenshots. P6, P7, and P8 previously had to write small Python programs on their own to analyze brain tissue visually.

9.3 Findings

Now, we describe the findings and lessons learned during the qualitative user study. These findings are based on the form participants filled out after each session and oral feedback they gave us during the evaluation.

Domain experts' interests are highly diverse

Therefore, they appreciate the customization functionalities of NeuroKit. For instance, some neurobiologists were only interested in the relationship between mitochondria and surrounding synapses. In contrast, others were solely interested in the synapses related to the mitochondrion's nerve cell. Furthermore, domain experts indicated that they want

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

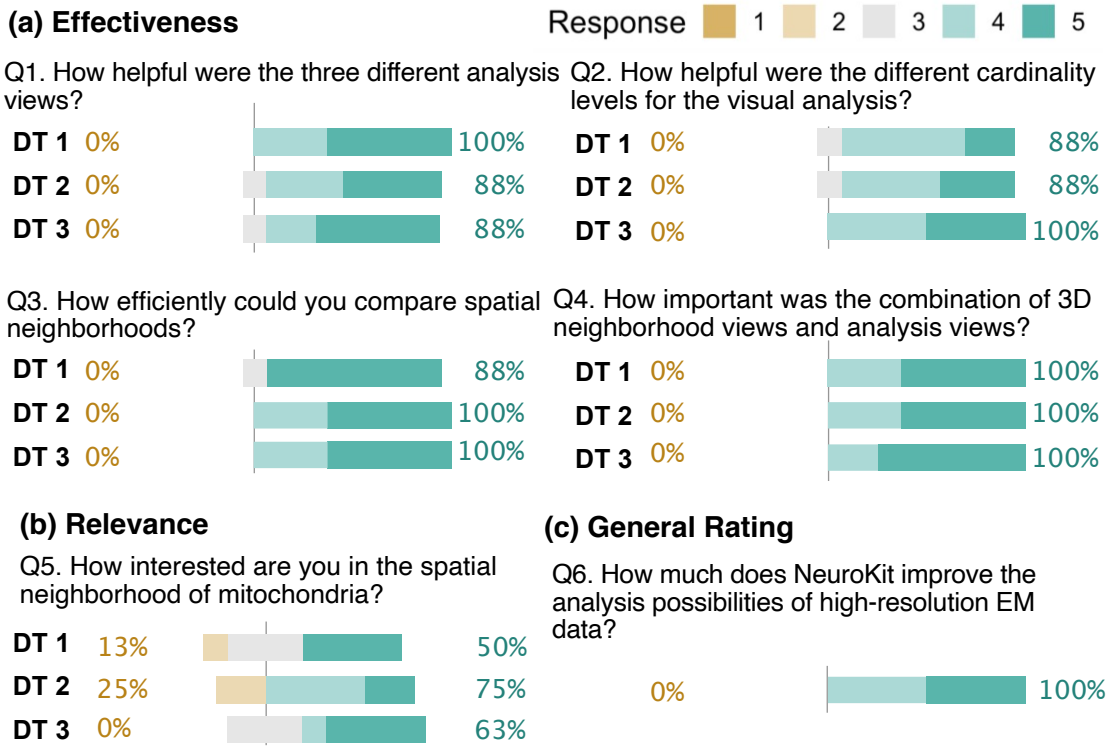


Figure 9.2: **NeuroKit user ratings.** We show negative responses (1, 2) on the left of each bar as a percentage number in beige. We show positive responses (4, 5) on the right side of each bar in teal. Results are based on the questionnaires shown in Fig. 9.3, 9.4, 9.5.

to collect new segmentation data to use NeuroKit with other structures of interest such as ribosomes or endoplasmatic reticula, or neuromuscular junctions.

Combining EM data with reconstructed data is essential

Domain experts highly appreciated overlaying the reconstructed 3D meshes with the actual EM image data. It allowed them to quickly verify whether segmentations were correct and enabled them to spot properties of brain tissue that had not been segmented. However, it was essential for them to toggle the overlay functionality interactively.

Experts use high-cardinality comparisons to spot trends

Most participants wanted to look at the high-cardinality visualization first to gain an overview of all neighborhoods in the data set and to spot outliers or specific neighborhoods of interest, which they then examined more closely in the neighborhood view.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Spine coverage of mitochondria is of high interest to all experts.

P6 mentioned that visualizing spine coverage using NeuroKit *"is of scientific significance"* and P2 stated *"this is awesome"*. Overall, experts were interested in analyzing spine coverage.

NeuroKit's user interface is easy and understandable

After a short introduction, it was generally straightforward for all participants to interpret and analyze all visualizations in NeuroKit. For instance, P3 mentioned that *"the visualizations are all really clear"*.

NeuroKit improves the state of the art

All participants indicated that NeuroKit is very useful and that the software has a high potential as a visual analysis tool for neuronal 3DEM data. P1, P2, P3, P6, and P7 plan to specify additional analysis scenarios for their data sets.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Analyzing the synaptic neighborhoods of mitochondria

In this section, we collect feedback on the effectiveness of the first user task, which investigates the spatial relationships between mitochondria and synapses.

* Erforderlich

How interested are you in the spatial relationship between mitochondria and synapses? *

	1	2	3	4	5	
not interested at all	○	○	○	○	○	very interested

How helpful were the three information visualizations to understand the spatial relationships between mitochondria and synapses? *

	1	2	3	4	5	
not helpful at all	○	○	○	○	○	very helpful

How helpful were the different levels of scale in the different views? *

	1	2	3	4	5	
not helpful at all	○	○	○	○	○	very helpful

How efficient could you compare multiple neighborhoods of mitochondria? *

	1	2	3	4	5	
not efficient at all	○	○	○	○	○	very efficient

How important was the combination of 3D and information visualizations for you to understand the relationship between mitochondria and synapses? *

	1	2	3	4	5	
not important at all	○	○	○	○	○	very important

Figure 9.3: Questionnaire for the domain task *Analyzing Synaptic Neighborhoods of Mitochondria*.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Analyzing mitochondria placement

In this section we collect feedback on the effectiveness of the first user task, which investigates the spatial relationships between mitochondria and their cell boundaries.

* Erforderlich

How interested are you in the spatial relationship between mitochondria and the cell boundary of the respective cell? *

	1	2	3	4	5	
not interested at all	○	○	○	○	○	very interested

How helpful were the three information visualizations to understand the spatial relationships between mitochondria and the cell boundary? *

	1	2	3	4	5	
not helpful at all	○	○	○	○	○	very helpful

How helpful were the different levels of scale in the different views? *

	1	2	3	4	5	
not helpful at all	○	○	○	○	○	very helpful

How efficient could you compare multiple neighborhoods of mitochondria? *

	1	2	3	4	5	
not efficient at all	○	○	○	○	○	very efficient

How important was the combination of 3D and information visualizations for you to understand the relationship between mitochondria and synapses? *

	1	2	3	4	5	
not important	○	○	○	○	○	very important

Figure 9.4: Questionnaire for the domain task *Analyzing Mitochondria Placement*.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Analyzing mitochondria spine coverage

In this section we collect feedback on the effectiveness of the first user task, which investigates the spatial relationships between mitochondria and dendritic spines.

How interested are you in the spatial relationship between mitochondria and dendritic spines? *

	1	2	3	4	5	
not interested at all	○	○	○	○	○	very interested

How helpful were the three information visualizations to understand the spatial relationships between mitochondria and dendritic spines? *

	1	2	3	4	5	
not helpful at all	○	○	○	○	○	very helpful

How helpful were the different levels of scale in the different views? *

	1	2	3	4	5	
not helpful at all	○	○	○	○	○	very helpful

How efficient could you compare multiple neighborhoods of mitochondria? *

	1	2	3	4	5	
not efficient at all	○	○	○	○	○	very efficient

How important was the combination of 3D and information visualizations for you to understand the relationship between mitochondria and dendritic spines? *

	1	2	3	4	5	
not important	○	○	○	○	○	very important

Figure 9.5: Questionnaire for the domain task *Analyzing Mitochondria Spine Coverage*.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

CHAPTER 10

Conclusions and Future Work

In this Chapter, we conclude the thesis by summarizing our work and discussing promising future research directions.

10.1 Summary

We presented NeuroKit, a visual tool to analyze and compare spatial neighborhoods in brain tissue. NeuroKit efficiently supports scientists in their data exploration and analysis process by an automatic guidance that adjusts the default user interface, visualization parameters, and interactions. Combining spatial neighborhood analysis with a scalable and customizable comparative visualization framework contributes to both the scientific field of neurobiology and data visualization. When creating domain-specific visualization tools, researchers often need the design to follow either a one-fits-it-all approach or an approach specifically designed around the needs of the domain experts. By creating NeuroKit as a customizable toolkit, we aim to get the best of both worlds by being general enough to target a wide range of experts and providing them enough detailed analysis capabilities to investigate their specific research interests.

Our user study found that NeuroKit meets domain experts' needs by supporting the analysis of different data cardinalities and combining detailed real-time 3D renderings with more abstract and quantitative analysis views. We support the highly diverse needs of neurobiologists by allowing specific customizations of NeuroKit.

10.2 Discussion

10.2.1 Additional features

During the user study, participants suggested a couple of additional features for NeuroKit. For instance, it would be interesting to extend NeuroKit with visualizations for analyzing

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

the texture of mitochondria in greater detail. P6 and P7 have classified mitochondria based on their texture in EM images into bright, dark, and combined mitochondria. They would be interested in relating these categorizations with our spatial neighborhood analysis. This would require further image-based analyses of our data sets and optimally larger data sets to investigate the differences of mitochondria in different brain regions. Additionally, overlaying the EM slice view with the segmentation data was requested by some of the users.

10.2.2 Limitations

Due to the high resolution of EM data, currently, our data sets do not include full reconstructions of nerve cells, but only small parts of them. Therefore, neurites and cell organelles are often cut off, limiting the quantitative analysis of those incomplete structures. Additionally, some pre-processing steps such as segmentation and mesh-feature extractions still require manual work. For instance, the dendritic spines and the cell body of all axons and dendrites were manually labeled in the surface meshes. This information is required to compute the visualizations to investigate the spine coverage of mitochondria. Manual data labeling and pre-processing make analyzing larger data sets quite complicated. In the future, we plan to automate data pre-processing further.

We offer a human-readable format that does not require programming and is well received for specifying analysis scenarios. It still requires manual editing and a rather technical description of analysis scenarios. In the future, we could further simplify the ADF specification by offering a visual interface to it.

10.3 Future Work

We want to prepare NeuroKit to be usable in the scientific community by providing detailed examples and documentation. Video tutorials will also help to make NeuroKit widely accessible. We suggest two primary research directions that could improve state of the art in visualization of connectomics data. First, applying the approaches of NeuroKit to much larger data sets with more diverse types of cell organelles is a promising direction. Second, we plan to incorporate a wider range of visualization methods and machine learning techniques to automatically learn properties of spatial neighborhoods in large scale data sets.

10.3.1 Larger Data Sets

In the future, we want to use NeuroKit to study larger data sets and therefore make the visual analysis more informative. We also plan to incorporate additional usage scenarios involving other cell organelles such as ribosomes, endoplasmatic reticula, and neuromuscular junctions once their EM segmentations and reconstructions are available. Also, we want to investigate how mitochondria cover boutons in axons, similar to mitochondria spine coverage.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

10.3.2 Additional Analysis & Visualization Methods

During the evaluation of NeuroKit, domain experts pointed out that knowledge about spatial neighborhoods of cell organelles such as mitochondria can help classify cell types. For instance, small mitochondria with proximity to the cell membrane most likely belong to axons. We plan to use these neighborhood characteristics to investigate to which extent they allow us to label neuronal structures automatically. We also plan to incorporate more topology-centered methods to classify neurites and their spatial neighborhoods.

10.4 Conclusion

We believe that our neighborhood analysis approach is generalizable to other areas, such as single-cell analyses or even exploring local neighborhoods in large city models. Our approach of supporting customized analysis scenarios could be helpful in many domain-specific visual analysis tools where users might have varying analysis goals. Finally, we believe that NeuroKit can advance neuroscience research, not just for the presented usage scenario of mitochondria analysis, but also for future data sets and future novel domain questions that have yet to be specified by neuroscientists.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

List of Figures

1.1 **Customized scalable comparison of spatial neighborhoods.** NeuroKit supports user-defined analysis scenarios (a,b) and automatically updates and adjusts visualizations and view parameters, to guide users in their specific analyses. Quantitative analysis views (c, d) enable scalable comparisons at different cardinality levels. Detailed 3D neighborhood views (e, f) depict the surroundings of structures of interest. Synapse domain task: (a) Synapse analysis scenario. (c) Scatterplot of hundreds of synapses from two different specimen. (e) Selected synapse (purple) and its neighborhood. Mitochondria domain task: (b) Mitochondria analysis scenario. (d) Distance-based abstraction of dendritic spines and their relation to mitochondria. (f) Mitochondria color-coded with their distance to the neurite membrane. 2

1.2 **NeuroKit design.** (a) Users can specify custom analysis scenarios in *analysis definition files*. Each user-defined *scenario* can consist of multiple *domain tasks* and incorporate customized visualizations and view parameters. (b) Our comparative visualization framework supports *low-cardinality*, *medium-cardinality*, or *high-cardinality* comparative analysis views (b, top row). Based on the number of structures in a comparison, NeuroKit automatically shows the corresponding view. To analyze complex 3D neighborhoods in more detail, we link quantitative views with 3D neighborhood views (b, bottom row). (c) Scientists can interactively refine the visual analysis to gain insight into their complex data sets. 5

2.1 **Different scales of visualizing brain connectivity.** NeuroKit focuses on a micro-scale analysis (c). Source: (a) Macro [GHK+08], (b) Meso [Seu12], (c) Micro [PKB+12] 7

2.2 **Scalar visualization of DTI data** using mean diffusivity and fractional anisotropy. Source: [VZKL06]. 8

2.3 **DTI based connectivity visualizations.** Superquadrics visualizing the diffusion tensor per voxel. By tracing streamlines through the field of principal eigenvectors a full-brain tractography is derived. Source: [Kin04, BBP+05]. 10

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

2.4 **Micro-scale brain connectivity visualization** with Neurolines [AABS+14]. 3D rendering of two neurons forming a synapse. Skeletons of these two neurons abstracted using a subway metaphor. Links between the skeletons illustrate synapses. 12

2.5 **Four different visual comparison approaches.** Each column shows a specific visual comparison approach (from left to right: Juxtaposition, Superimposition, Interchangeable, and Explicit Encoding). The top row indicates the comparison of two entities while the bottom row shows how the respective approach applies to comparing N entities. Source: [KCK17] 14

2.6 **Examples of visualizations generated with Vega-Lite.** (a) JSON specification of a line chart and the resulting visualization. (b) JSON specifiction of a scatter plot and the resulting visualization. Source: [SMWH17] . . . 17

3.1 **Structure of neurons and cell organelles.** Nerve signals are transmitted from axons to dendrites via synapses. Neurons contain several different cell organelles, such as mitochondria or ribosomes. Mitochondria are responsible for providing energy to the neuron while ribosomes synthesize proteins. Source: Johanna Beyer 20

3.2 **Steps of the connectomics workflow.** Source: [Moh17] 21

3.3 **Slice of brain tissue imaged with electron microscopy** at a resolution of 5×5 microns. Data provided by Graham Knott [KMWL08]. 22

3.4 **Segmentation masks of EM connectomics data.** Each color relates to a neuronal structure. Source: Brian Matejek [MHL+17] 23

3.5 **Six different data sets acquired by our collaborators [KMWL08].** The first column shows a 3D rendering of the respective data set and the connectivity of nerve cells as a node-link diagram. The data has a cubic geometry since it shows a small subvolume of the mouse brain. The second column shows information data of the data set and the third column holds information about the segmented structures. Source: [ACA+19] 25

4.1 Overview of the design study methodology approach by Sedlmair et al. [SMM12] 28

4.2 **Connectivity vs. neighborhood analysis.** Connectivity analysis of neuronal structures (a) compared to spatial neighborhood analysis (b). Our collaborators are specifically interested in the latter. 30

5.1 **NeuroKit workflow and automatic user guidance.** NeuroKit guides and supports users throughout the analysis workflow to minimize mental load and support specific analysis tasks. We show user actions in purple, and automatic guidance actions in gray. A user can customize the visual analysis by specifying a user-defined analysis scenario. Based on the analysis scenario, we automatically adjust the visual user interface and propose the most appropriate visualization and comparison parameters. 36

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

6.1 **NeuroKit user interface.** (a) List view for selecting structures of interest. (b) 3D overview of all selected structures of interest. (c) Analysis view for detailed comparisons of the selected local neighborhoods. (d) Detailed 3D neighborhood views. (e) Panel for user-defined analysis scenario, showing suggested comparison modes and custom UI elements. 41

6.2 **Detailed three-dimensional neighborhood view.** The dendrite (green) developed spines that form synapses (purple), and an enclosed mitochondrion (surface color-coded structure). Orange/reddish color on the surface indicates close proximity between the mitochondrion and the dendrite's cell membrane, while blue denotes a greater distance. The 3D rendering is overlayed with a 2D slice of the original EM data. 43

7.1 **Volume boundary handling for distance computations.** Neurites (orange, green) and cell organelles such as mitochondria (red) are often cut off by the volumes cutting plane (outer box). To prevent a distance distribution from being incorrectly calculated, we introduce a margin (space between boxes) that is not considered for distance distributions. 46

7.2 **Volume computation without transforming the meshes into a volumetric data representation.** Instead, the mesh volume can be calculated by the volume of the tetrahedron each triangle forms with the origin of the coordinate space. Source: [ZC01] 48

7.3 The real-time rendering pipeline. Source: [Wim20]. 50

7.4 3D rendering using per-pixel linked lists for order independent transparency (OIT). 51

8.1 **Two different synaptic neighborhoods of mitochondria.** (a) all synapses that are within 1 micron from the mitochondrion and (b) synapses connected to the mitochondrion's dendrite. 62

8.2 **Analysis views for the domain task *Synaptic Neighborhood of Mitochondria.*** (a) - (c): low-, medium-, high-cardinality comparison. (a) distance tree, (b) grouped bar chart, (c) distance matrix. 64

8.3 **Two mitochondria placed differently in their cells.** The axonal mitochondrion (a) is much closer to the cell membrane, than the dendritic mitochondrion (b). A red color on the mitochondrions surface indicates close proximity while blue represents a greater distance. 65

8.4 **Analysis views for the domain task *Mitochondria Placement.*** (a)-(c): low-, medium-, high-cardinality comparison. (a) cumulative histogram, (b) violin plot, (c) scatter plot. 67

8.5 **Spine coverage.** All spines (tiny branches) are covered by a mitochondrion. 68

8.6 **Analysis views for the domain task *Mitochondric Spine Coverage.*** (a) - (c): low-, medium-, high-cardinality comparison. (a) dendrite abstraction level 1, (b) dendrite abstraction level 2, (c) scatter plot. 70

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

8.7 **Screenshot of NeuroKit s' user interface** showing the domain task *Mitochondric Spine Coverage* in the medium cardinality visualization (dendrite abstraction level 2). 71
8.8 **Screenshot of NeuroKit s' user interface** showing the domain task *Mitochondria Placement* in the low cardinality visualization (cumulative histogram). 72

9.1 **Comparing synapse locations in two different mouse data sets.** Grouped bars show the distances between a mitochondrion and its neighboring synapses. Hovering over a bar highlights the corresponding synapse in the 3D neighborhood view. The young mouse (purple, a) exhibits *more* and *closer* synapses than the adult mouse (green, b), indicating synaptic pruning over time. . 76
9.2 **NeuroKit user ratings.** We show negative responses (1, 2) on the left of each bar as a percentage number in beige. We show positive responses (4, 5) on the right side of each bar in teal. Results are based on the questionnaires shown in Fig. 9.3, 9.4, 9.5. 78
9.3 Questionnaire for the domain task *Analyzing Synaptic Neighborhoods of Mitochondria*. 80
9.4 Questionnaire for the domain task *Analyzing Mitochondria Placement*. . . . 81
9.5 Questionnaire for the domain task *Analyzing Mitochondria Spine Coverage*. 82

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Bibliography

[AABH+16] A. Al-Awami, J. Beyer, D. Haehn, N. Kasthuri, J. W. Lichtman, H. Pfister, and M. Hadwiger. NeuroBlocks – Visual Tracking of Segmentation and Proofreading for Large Connectomics Projects. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):738–746, 2016.

[AABS+14] A. Al-Awami, J. Beyer, H. Strobelt, N. Kasthuri, J. W. Lichtman, H. Pfister, and M. Hadwiger. NeuroLines: A Subway Map Metaphor for Visualizing Nanoscale Neuronal Connectivity. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2369–2378, 2014.

[ABHR+13] B. Alper, B. Bach, N. Henry Riche, T. Isenberg, and J.D. Fekete. Weighted graph comparison techniques for brain connectivity analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 483–492, 2013.

[ACA+19] M. Agus, C. Calì, A. Al-Awami, E. Gobbetti, P. Magistretti, and M. Hadwiger. Interactive Volumetric Visual Analysis of Glycogen-derived Energy Absorption in Nanometric Brain Structures. *Computer Graphics Forum*, 38(3):427–439, 2019.

[ACG+09] F. Azevedo, L. Carvalho, L. Grinberg, J. Farfel, R. Ferretti, R. Leite, W. Filho, R. Lent, and S. Herculano-Houzel. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *Journal of Comparative Neurology*, 513(5):532–541, 2009.

[AHKGF11] R. Arias-Hernandez, L. T. Kaastra, T. M. Green, and B. Fisher. Pair Analytics: Capturing Reasoning Processes in Collaborative Visual Analytics. In *Hawaii International Conference on System Sciences*, pages 1–10, 2011.

[ATW20] P. Alliez, S. Tayeb, and C. Wormser. 3D Fast Intersection and Distance Computation. In *CGAL User and Reference Manual*. CGAL Editorial Board, 5.1 edition, 2020.

[BAAK+13] J. Beyer, A. Al-Awami, N. Kasthuri, J. W. Lichtman, H. Pfister, and M. Hadwiger. ConnectomeExplorer: Query-Guided Visual Analysis of

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Large Volumetric Neuroscience Data. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2868–2877, 2013.

[BBP+05] J. Blaas, C.P. Botha, B. Peters, F.M. Vos, and F.H. Post. Fast and Reproducible Fiber Bundle Selection in DTI Visualization. In *VIS 05. IEEE Visualization, 2005.*, pages 59–64, Minneapolis, MN, USA, 2005. IEEE.

[BH09] M. Bostock and J. Heer. Protovis: A Graphical Toolkit for Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1121–1128, 2009.

[BHAA+13] J. Beyer, M. Hadwiger, A. Al-Awami, Won-Ki J., N. Kasthuri, J. W. Lichtman, and H. Pfister. Exploring the Connectome: Petascale Volume Visualization of Microscopy Data Streams. *IEEE Computer Graphics and Applications*, 33(4):50–61, 2013.

[BOH11] M. Bostock, V. Ogievetsky, and J. Heer. D3 Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309, 2011.

[BPP+00] P. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, 44(4):625–632, 2000.

[BSL18] D. Berger, H. S. Seung, and Jeff W. Lichtman. VAST (Volume Annotation and Segmentation Tool): Efficient Manual and Semi-Automatic Labeling of Large 3D Image Stacks. *Frontiers in Neural Circuits*, 12, 2018.

[CBB+16] C. Calì, J. Baghabra, D. Boges, G. Holst, A. Kreshuk, F. Hamprecht, M. Srinivasan, H. Lehväslaiho, and P.J. Magistretti. Three-dimensional immersive virtual reality for studying cellular compartments in 3D models from EM preparations of neural tissues. *Journal of Comparative Neurology*, 524(1):23–38, 2016.

[CSS+12] A. Cardona, S. Saalfeld, J. Schindelin, I. Arganda-Carreras, S. Preibisch, M. Longair, P. Tomancak, V. Hartenstein, and R.J. Douglas. TrakEM2 Software for Neural Circuit Reconstruction. *PLoS One*, 7(6):e38011, 2012.

[CWB+18] C. Calì, M. Wawrzyniak, C. Becker, B. Maco, M. Cantoni, A. Jorstad, B. Nigro, F. Grillo, V. De Paola, P. Fua, and others. The effects of aging on neuropil structure in mouse somatosensory cortex—A 3D electron microscopy analysis of layer 1. *PLoS One*, 13(7):e0198131, 2018.

[Eul45] L. Euler. *Institutionum calculi integralis*, volume 4. impensis Academiae imperialis scientiarum, 1845.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[FGBHOdS05] R. Fernandez-Gonzalez, M. Barcellos-Hoff, and C. Ortiz-de Solorzano. A tool for the quantitative spatial analysis of complex cellular systems. *IEEE Transactions on Image Processing*, 14(9):1300–1313, 2005.

[FGK+00] A. Fabri, G. Giezeman, L. Kettner, S. Schirra, and S. Schönherr. On the design of CGAL a computational geometry algorithms library. *Software: Practice and Experience*, 30(11):1167–1202, 2000.

[GAW+11] M. Gleicher, D. Albers, R. Walker, I. Jusufi, C.D. Hansen, and J.C. Roberts. Visual comparison for information visualization. *Information Visualization*, 10(4):289–309, 2011.

[GG90] J. Glaser and E. Glaser. Neuron imaging with Neurolucida—a PC-based system for image combining microscopy. *Computerized Medical Imaging and Graphics*, 14(5):307–317, 1990.

[GGW+00] W. Gan, J. Grutzendler, W. Wong, R. Wong, and J.W. Lichtman. Multicolor "DiOlistic" labeling of the nervous system using lipophilic dye combinations. *Neuron*, 27(2):219–225, 2000. Publisher: Elsevier.

[GHK+08] X. Gigandet, P. Hagmann, M. Kurant, L. Cammoun, R. Meuli, and J.P. Thiran. Estimating the Confidence Level of White Matter Connections Obtained with MRI Tractography. *PLOS ONE*, 3(12):e4006, December 2008.

[GWB+21] F. Gonda, X. Wang, J. Beyer, M. Hadwiger, J.W. Lichtman, and H. Pfister. VICE: Visual Identification and Correction of Neural Circuit Errors. *Computer Graphics Forum (Proceedings Eurographics/IEEE Symposium on Visualization, Eurovis 2021*, page to appear, 2021.

[GWP21] F. Gonda, D. Wei, and H. Pfister. Consistent Recurrent Neural Networks for 3D Neuron Segmentation. *arXiv preprint arXiv:2102.01021*, 2021.

[HAAB+17] M. Hadwiger, A.K. Al-Awami, J. Beyer, M. Agus, and H. Pfister. SparseLeap: Efficient empty space skipping for large-scale volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):974–983, 2017.

[HBJP12] M. Hadwiger, J. Beyer, W.-K Jeong, and H. Pfister. Interactive Volume Exploration of Petascale Microscopy Data Streams Using a Visualization-Driven Virtual Memory Approach. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2285–2294, 2012.

[HHM+17] D. Haehn, J. Hoffer, B. Matejek, A. Suissa-Peleg, A. Al-Awami, L. Kamentsky, F. Gonda, E. Meng, W. Zhang, R. Schalek, A. Wilson, T. Parag, J. Beyer, V. Kaynig, T.R. Jones, J. Tompkin, M. Hadwiger, J.W. Lichtman, and H. Pfister. Scalable Interactive Visualization for Connectomics. *Informatics*, 4(3):29, 2017.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[HKT+18] D. Haehn, V. Kaynig, J. Tompkin, J.W. Lichtman, and H. Pfister. Guided proofreading of automatic segmentations for connectomics. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9319–9328, 2018.

[HLL97] L. Hesselink, Y. Levy, and Y. Lavin. The topology of symmetric, second-order 3D tensor fields. *IEEE Transactions on Visualization and Computer Graphics*, 3(1):1–11, March 1997.

[HSKK01] M. Hilaga, Y. Shinagawa, T. Kohmura, and T.L. Kunii. Topology matching for fully automatic similarity estimation of 3D shapes. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '01, pages 203–212, August 2001.

[JNC+15] A. Jorstad, B. Nigro, C. Cali, M. Wawrzyniak, P. Fua, and G. Knott. NeuroMorph: A Toolset for the Morphometric Analysis and Visualization of 3D Models Derived from Electron Microscopy Image Stacks. *Neuroinformatics*, 13(1):83–92, 2015.

[KBJ+20] R. Krueger, J. Beyer, W.-D. Jang, N. W. Kim, A. Sokolov, P. K. Sorger, and H. Pfister. Facetto: Combining Unsupervised and Supervised Learning for Hierarchical Phenotype Analysis in Multi-Channel Image Data. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):227–237, 2020.

[KCK17] K. Kim, J. Carlis, and D. Keefe. Comparison techniques utilized in spatial 3D and 4D data visualizations: A survey and future directions. *Computers & Graphics*, 67:138–147, 2017. Publisher: Elsevier.

[KH13] J. Kehrer and H. Hauser. Visualization and Visual Analysis of Multifaceted Scientific Data: A Survey. *IEEE Transactions on Visualization and Computer Graphics*, 19(3):495–513, March 2013.

[Kin04] Gordon Kindlmann. Superquadric Tensor Glyphs. In *Proceedings of the Sixth Joint Eurographics - IEEE TCVG Conference on Visualization*, VISSYM'04, pages 147–154, 2004.

[KMWL08] G. Knott, H. Marchman, D. Wall, and B. Lich. Serial Section Scanning Electron Microscopy of Adult Brain Tissue Using Focused Ion Beam Milling. *Journal of Neuroscience*, 28(12):2959–2964, 2008.

[KWH00] G. Kindlmann, D. Weinstein, and D. Hart. Strategies for direct volume rendering of diffusion tensor fields. *IEEE Transactions on Visualization and Computer Graphics*, 6(2):124–138, June 2000.

[LAK+98] D.H. Laidlaw, E.T. Ahrens, D. Kremers, M.J. Avalos, R.E. Jacobs, and C. Readhead. Visualizing diffusion tensor images of the mouse spinal

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

cord. In *Proceedings Visualization '98*, pages 127–134, Research Triangle Park, NC, USA, 1998. IEEE.

[LBK+18] F. Lekschas, B. Bach, P. Kerpedjiev, N. Gehlenborg, and H. Pfister. HiPiler: Visual Exploration of Large Genome Interaction Matrices with Interactive Small Multiples. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):522–531, 2018.

[LD11] J. W. Lichtman and W. Denk. The Big and the Small: Challenges of Imaging the Brain's Circuits. *Science*, 334(6056):618–623, November 2011.

[LLS08] J.W. Lichtman, J. Livet, and J.R. Sanes. A technicolour approach to the connectome. *Nature Reviews Neuroscience*, 9(6):417–422, June 2008.

[LW19] Z. Lin and D. Wei. PyTorch Connectomics, 2019. https://github.com/zudi-lin/pytorch_connectomics.

[MAAB+18] H. Mohammed, A. Al-Awami, J. Beyer, C. Calì, P. Magistretti, H. Pfister, and M. Hadwiger. Abstractocyte: A Visual Tool for Exploring Nanoscale Astroglial Cells. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):853–861, 2018.

[MGT+03] T. Munzner, F. Guimbretière, S. Tasiran, L. Zhang, and Y. Zhou. TreeJuxtaposer: scalable tree comparison using Focus+Context with guaranteed visibility. *ACM Transactions on Graphics*, 22(3):453–462, 2003.

[MHL+17] B. Matejek, D. Haehn, F. Lekschas, M. Mitzenmacher, and H. Pfister. Compresso: Efficient Compression of Segmentation Data for Connectomics. In M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D.L. Collins, and S. Duchesne, editors, *Medical Image Computing and Computer Assisted Intervention-MICCAI 2017*, Lecture Notes in Computer Science, pages 781–788, Cham, 2017. Springer International Publishing.

[MMH+13] A. Maries, N. Mays, M. Hunt, K. F. Wong, W. Layton, R. Boudreau, C. Rosano, and G. E. Marai. GRACE: A Visual Comparison Framework for Integrated Spatial and Non-Spatial Geriatric Data. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2916–2925, 2013.

[Moh17] H. Mohammed. Abstractocyte: A Visual Tool for Exploring Nanoscale Astroglial Cells. Master's thesis, King Abdullah University of Science and Technology, 2017.

[MvZ02] S. Mori and P. van Zijl. Fiber tracking: principles and strategies – a technical review. *NMR in Biomedicine*, 15(7-8):468–480, 2002.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[NDBB17] H. Nhu, Rafael Arrojo E. Drigo, Per-Olof Berggren, and Thomas Boudier. A novel toolbox to investigate tissue spatial organization applied to the study of the islets of Langerhans. *Scientific reports*, 7(1):1–12, 2017.

[PB13] B. Preim and C.P. Botha. *Visual computing for medicine: theory, algorithms, and applications*. Newnes, 2013.

[Pet15] J.B. Pettit. *Spatial analysis of complex biological tissues from single cell gene expression data*. PhD Thesis, University of Cambridge, 2015.

[PKB+12] H. Pfister, V. Kaynig, C.P. Botha, S. Bruckner, V. Dercksen, H.C. Hege, and J. Roerdink. Visualization in Connectomics. *arXiv:1206.1428 [cs, q-bio]*, 2012.

[PP95] H.G. Pagendarm and F. Post. *Comparative visualization: approaches and examples*. Delft University of Technology, 1995.

[Rai19] R. Raidou. Lecture on Segmentation & Registration, November 2019. https://www.cg.tuwien.ac.at/courses/MedVis/VU/2019W.

[Run95] C. Runge. Über die numerische Auflösung von Differentialgleichungen. *Mathematische Annalen*, 46(2):167–178, 1895. Publisher: Springer.

[SBS+13] J. Sorger, K. Bühler, F. Schulze, T. Liu, and B. Dickson. neuroMAP — Interactive graph-visualization of the fruit fly's neural circuit. In *IEEE Symposium on Biological Data Visualization (BioVis)*, pages 73–80, 2013.

[SDL03] Z. Song, C. Demiralp, and D.H. Laidlaw. Visualizing diffusion tensor mr images using streamtubes and streamsurfaces. *IEEE Transactions on Visualization and Computer Graphics*, 9(4):454–462, October 2003.

[Seu12] H.S. Seung. *Connectome: How the brain's wiring makes us who we are*. HMH, 2012.

[SFG+20] C. Stoltzfus, J. Filipek, B. Gern, B. Olin, J. Leal, Y. Wu, M. Lyons-Cohen, J. Huang, C. Paz-Stoltzfus, C. Plumlee, and others. CytoMAP: a spatial analysis toolbox reveals features of myeloid cell organization in lymphoid tissues. *Cell reports*, 31(3):107523, 2020.

[SH14] A. Satyanarayan and J. Heer. Lyra: An Interactive Visualization Design Environment. *Computer Graphics Forum*, 33(3):351–360, 2014.

[SMM12] M. Sedlmair, M. Meyer, and T. Munzner. Design Study Methodology: Reflections from the Trenches and the Stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440, 2012.

[SMWH17] A. Satyanarayan, D. Moritz, K. Wongsuphasawat, and J. Heer. Vega-Lite: A Grammar of Interactive Graphics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):341–350, January 2017.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[SPA+14] J. Schmidt, R. Preiner, T. Auzinger, M. Wimmer, E. Gröller, and S. Bruckner. YMCA — Your mesh comparison application. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 153–162, 2014.

[SRHH16] A. Satyanarayan, R. Russell, J. Hoffswell, and J. Heer. Reactive Vega: A Streaming Dataflow Architecture for Declarative Interactive Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):659–668, 2016.

[SSKH11] C. Sommer, C. Straehle, U. Köthe, and F. A. Hamprecht. Ilastik: Interactive learning and segmentation toolkit. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 230–233, 2011.

[Stu08] R. Stufflebeam. Neurons, synapses, action potentials, and neurotransmission. *Consortium on Cognitive Science Instruction*, 2008.

[SW06] D. Sweeney and B. Williamson. *Biology: Exploring Life: Laboratory Manual*. Pearson Education, Incorporated, 2006.

[SWC+08] D. Schneider, A. Wiebel, H. Carr, M. Hlawitschka, and G. Scheuermann. Interactive Comparison of Scalar Fields Based on Largest Contours with Applications to Flow Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1475–1482, 2008.

[SWH14] A. Satyanarayan, K. Wongsuphasawat, and J. Heer. Declarative interaction design for data visualization. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, UIST '14, pages 669–678, 2014.

[SWKP18] S.D. Shank, S. Weaver, and S. Kosakovsky Pond. phylotree.js - a JavaScript library for application development and interactive data visualization in phylogenetics. *BMC Bioinformatics*, 19(1):276, July 2018.

[Tro21] J. Troidl. *NeuroKit*. 2021. https://github.com/jakobtroidl/NeuroKit.

[VCDR+20] E. Vezzoli, C. Calì, M. De Roo, L. Ponzoni, E. Sogne, N. Gagnon, M. Francolini, D. Braida, M. Sala, D. Muller, A. Falqui, and P.J. Magistretti. Ultrastructural Evidence for a Role of Astrocytes and Glycogen-Derived Lactate in Learning-Dependent Synaptic Stabilization. *Cerebral Cortex*, 30(4):2114–2127, 2020.

[VZKL06] A. Vilanova, S. Zhang, G. Kindlmann, and D. Laidlaw. An Introduction to Visualization of Diffusion Tensor Imaging and Its Applications. In J. Weickert and H. Hagen, editors, *Visualization and Processing of Tensor Fields*, pages 121–153. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[Wim20] M. Wimmer. Real-Time Rendering Lecture 2: The Graphics Pipeline, 2020. https://www.cg.tuwien.ac.at/courses/Realtime/VU/2020W.

[WLFB+20] D. Wei, Z. Lin, D. Franco-Barranco, N. Wendt, X. Liu, W. Yin, X. Huang, A. Gupta, W.D. Jang, X. Wang, and others. MitoEM Dataset: Large-Scale 3D Mitochondria Instance Segmentation from EM Images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 66–76. Springer, 2020.

[Wol11] D. Wolff. *OpenGL 4.0 shading language cookbook*. Packt Publishing Ltd, 2011.

[ZC01] C. Zhang and T. Chen. Efficient feature extraction for 2D/3D objects in mesh representation. In *Proceedings 2001 International Conference on Image Processing*, volume 3, pages 935–938. IEEE, 2001.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN