

浙江大学

本科毕业设计报告



题目： 在线交互式神经元重建的服务器系统

姓 名： 曲衡 3130000569

学 号： 曲衡 3130000569

指导教师： 曲衡 3130000569

专 业： 2013 级 计算机科学与技术

学 院： 计算机科学与技术学院

A Dissertation Submitted to Zhejiang
University for the Degree of Bachelor of
Engineering



TITLE: High-performance Server System for Online Interactive N

Author: Heng Qu 3130000569

StudentID: Heng Qu 3130000569

Mentor: Heng Qu 3130000569

College: College of Computer Science and Technology

Submitted Date:

浙江大学本科生毕业论文(设计)诚信承诺书

1. 本人郑重地承诺所呈交的毕业论文(设计),是在指导教师的指导下严格按照学校和学院有关规定完成的。
2. 本人在毕业论文(设计)中引用他人的观点和参考资料均加以注释和说明。
3. 本人承诺在毕业论文(设计)选题和研究内容过程中没有抄袭他人研究成果和伪造相关数据等行为。
4. 在毕业论文(设计)中对侵犯任何方面知识产权的行为,由本人承担相应的法律责任。

毕业论文(设计)作者签名:

_____年 ____月 ____日

摘 要

原始神经图像信息的神经元追踪与数字重建是神经科学界热门方向。神经元的形态反应出它的功能,相同结构的神经元通常具有类似的功能。神经科学家通过结构脑图谱的重建,可以反推大脑是如何运作的,对理解智慧的产生有重要的帮助。由于神经元拓扑结构的复杂性,在一些自动化重建结果的细节上仍然需要研究人员对数字重建的结果进行人工纠正和修改,以确保数字重建工作的准确性。另外研究人员需要对数字重建结果进行编辑,比如添加或删除一些网络分支等。现有的用于原始神经图像信息的神经元追踪和数字重建软件大多是运行在单机之上,无法满足多用户协同编辑与修改的需求,也不利于结构脑图谱的交流。计算机性能和网络速度的提升使得在线实时编辑神经网络结构成为了可能。在这样的背景下,设计并实现了在线多用户的神经元网络结构编辑分享平台,利用互联网便于数据共享与交流的特点,使得神经科学家可以便捷地进行异地,多用户协同编辑神经元网络结构,并能分享完成重建的结构脑图谱,探索神经元结构下的奥秘。项目中使用了 DVID 数据库储存神经元信息,PostgreSQL 储存用户信息并使用 Node.js 与 Express 完成了网络服务器的搭建,为前端可视化操作提供了有力的支持。实验与测试报告表明,该平台足以支撑至少数千名用户同时编辑的需求,并能在毫秒级别的时间内给出做出相应,达到了实时操作的要求,为神经科学家在神经元追踪以及数字重建方面提供了便利。

关键词: 生物图像信息神经元重建 DVID 实时编辑平台

Abstract

Neuron tracing and digital reconstruction from original neural image information is a hot direction of neural science. The morphology of neurons reflects its function, the neurons with same morphology usually have similar function. By the reconstruction of brain connectome, neuroscientists can speculate how the brain works which is helpful to understand the intelligence. Because of the complexity of the topological structure of neurons, and in some details on the results of automation reconstruction researchers still need to manually correct and modify the results of digital reconstruction, in order to ensure the accuracy of the digital reconstruction. In addition, researchers need to edit the result of the digital reconstruction such as add or delete some branches, etc. Existing neurons tracking for original neural image and digital reconstruction softwares are mostly run on stand-alone, unable to meet the requirements of multi-user collaborative editing and modification, but also not conducive to the exchange of brain connectome. Improvements in computer performance and network speed make it possible for online real-time editing of neural network structures. In such a background, We have designed and implemented a neural network structure of online multi-user edit sharing platform, using the Internet which can facilitate data sharing and exchange can easily carry out off-site, edit neural network structure with multi-user cooperation, and share brain connectome, explore the mystery of neuronal structure. Project uses DVID database to store neural information, PostgreSQL to store user information and Node.js, Express to build web server which provides the powerful support for the front-end visualization. Experiments and test reports show that the platform is sufficient to support the needs of at least thousands of users editing at the same time, response in milliseconds, meets the real-time operation requirements which provides convenience for neurologists in neuronal tracking and digital reconstruction.

Keywords: bioimage informatics; neuron tracing; DVID; real-time editing platform

目 录

摘要 I

Abstract II

目录 III

第 1 章 概论 1

 1.1 交互式神经元重建系统的背景与意义 1

 1.2 现有交互式神经元重建系统的特点与问题 3

 1.3 论文结构 5

第 2 章 整体架构与技术选型 7

 2.1 整体架构 7

 2.2 技术选型 9

第 3 章 平台搭建与性能调优 11

 3.1 数据表设计 11

第 4 章 项目成果 13

 4.1 第一节 13

参考文献 14

致谢 15

附录 16

第 1 章 概论

1.1 交互式神经元重建系统的背景与意义

原始神经元图像信息的神经元追踪和数字重建是神经科学界热门方向。神经元的形态反应出它的功能,相同功能的神经元通常具有类似的功能。神经科学家通过结构脑图谱的重建,可以反推大脑是如何运作,对理解智慧的产生有重要的帮助。十九世纪以来,神经科学家们开始推测记忆,甚至个性与智力都储存大脑神经元之间的连接里。图 1.1 展示了秀丽隐杆线虫的神经结构的神经结构,图中每一个节点均代表一个神经元,每一条线代表一个连接。它仅仅由 300 个神经元组成,之间的连接也仅有 7000 个。

White, John G 与 Southgate 等人在 1986 年时已经利用一系列局部原始电子显微照片对秀丽隐杆线虫的神经系统的进行了完整重建^[7]。经过了 30 多年的发展, Yunkyu Sohn, Myung-Kyu Choi 与 Yong-Yeol Ahn 等人于 2011 年利用基于模块化的群态检测算法发现秀丽隐杆线虫中有 5 个解剖簇及其对应的实验可识别功能电路,进一步揭示了生物电路如何产生更高阶的复杂行为^[6]。即使如此,由于神经网络复杂的拓扑结构,神经科学家们仍旧未能充分探索通过突触交织的神经网络结构。而人类大脑由一千亿个神经元组成,神经元之间的连接的数量又是神经元数量的一万倍,比秀丽隐杆线虫的神经结构要复杂的多。设计并实现出自动神经元重建算法便成了探索神经结构的重要步骤之一。

Druckmann, Shaul 与 Feng 等人开发的神经元重建算法提供了准确的中线,直径,表面,体积和分支点位置,支持沿着神经元表面分析标记过的分子分布,还可以直接导出到建模软件。图 1.2 展示了这种神经元重建算法的样例结果。Brown, Kerry M 与 Barrionuevo 等人收集了来自不同动物,脑区,神经元类型和可视化方法的六个数据集,为自动化软件所需的测试提供了基准,提高了重建的质量,同时最大限度地减少了人工的参与,极大的促进了神经元重建领域的发展^[1]。

由于神经元拓扑结构的复杂性,在一些自动化重建结果的细节上仍然需要研究

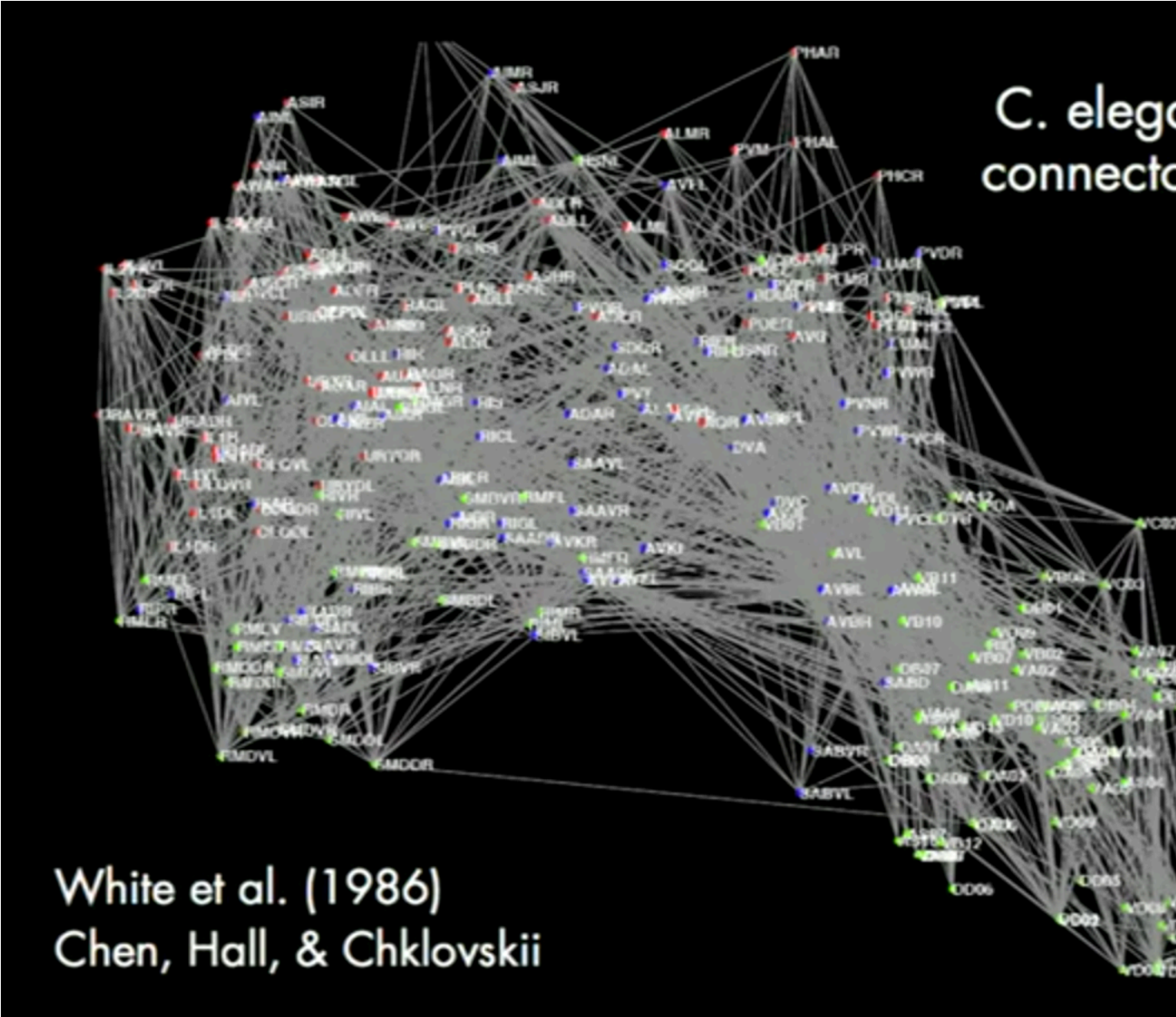


图 1.1 秀丽隐杆线虫的神经结构

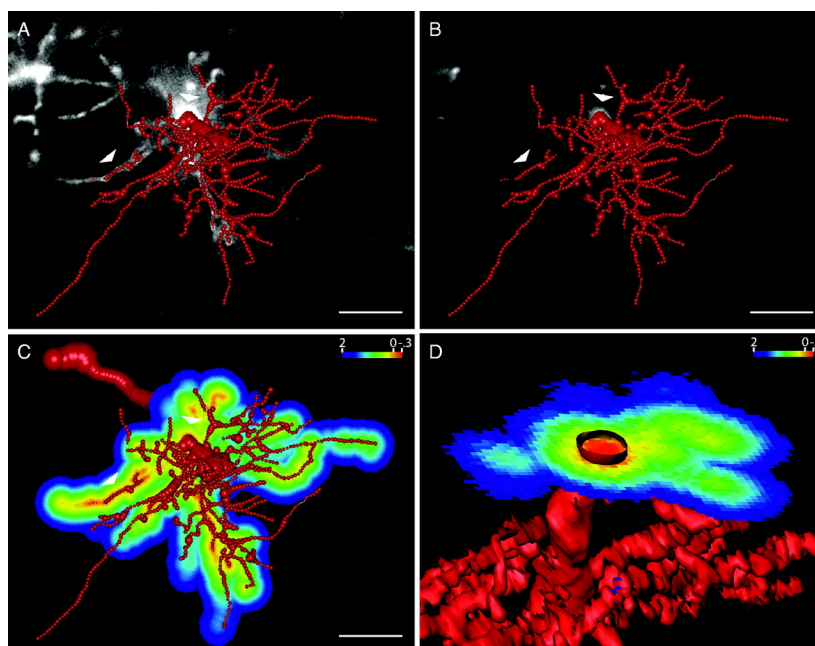


图 1.2 Druckmann 等人的神经元重建算法的样例结果

人员对数字重建的结果进行人工纠正和修改,以确保数字重建工作的准确性。另外研究人员需要对数字重建结果进行编辑,比如添加或删除一些网络分支等。为了便于研究人员进一步研究神经结构,探索智能产生的原因,这就需要在神经元自动重建算法的基础上建立交互式神经元重建系统。

1.2 现有交互式神经元重建系统的特点与问题

1.2.1 FARSIGHT

FARSIGHT 的软件界面如图 1.3 所示,图中正在编辑的是一小段神经结构,展示了用户点击位置,根节点方向,以及分支节点。

FARSIGHT 的设计目标是重建结果细节丰富,可以快速识别重建结果的错误并能迅速纠正。FARSIGHT 利用基于模式分析辅助集群编辑(PACE)的思想,根据对自动跟踪结果的定量测量和多变量模式分析工具的分析结果,发现了常见类型的重建错误,提高了纠正重建结果的效率^[3]。图 1.4 展示了 FARSIGHT 导出的较大规模的神经结构重建结果。

FARSIGHT 的缺点在于,它专注于半自动重建,对于常见的错误修改效率确实

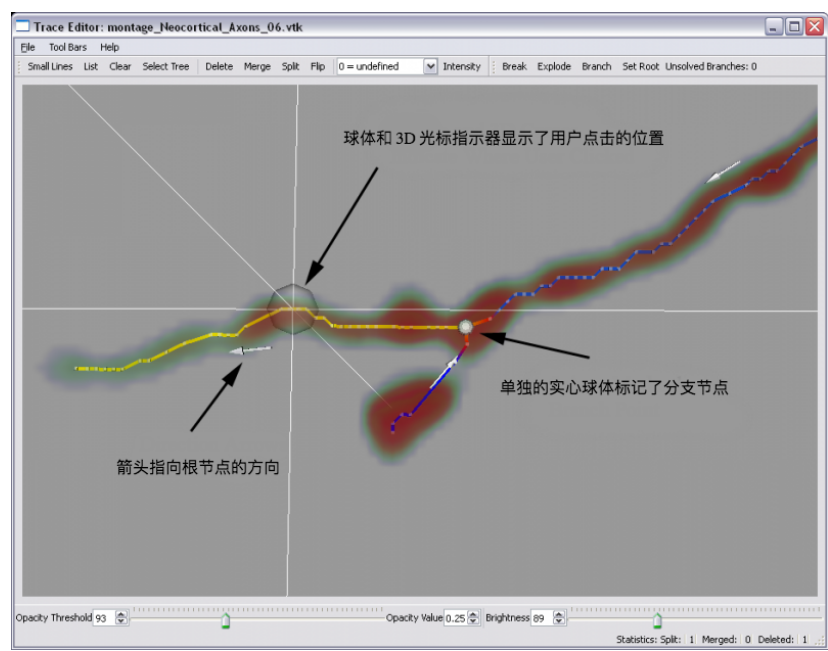


图 1.3 FARSIGHT 软件运行界面

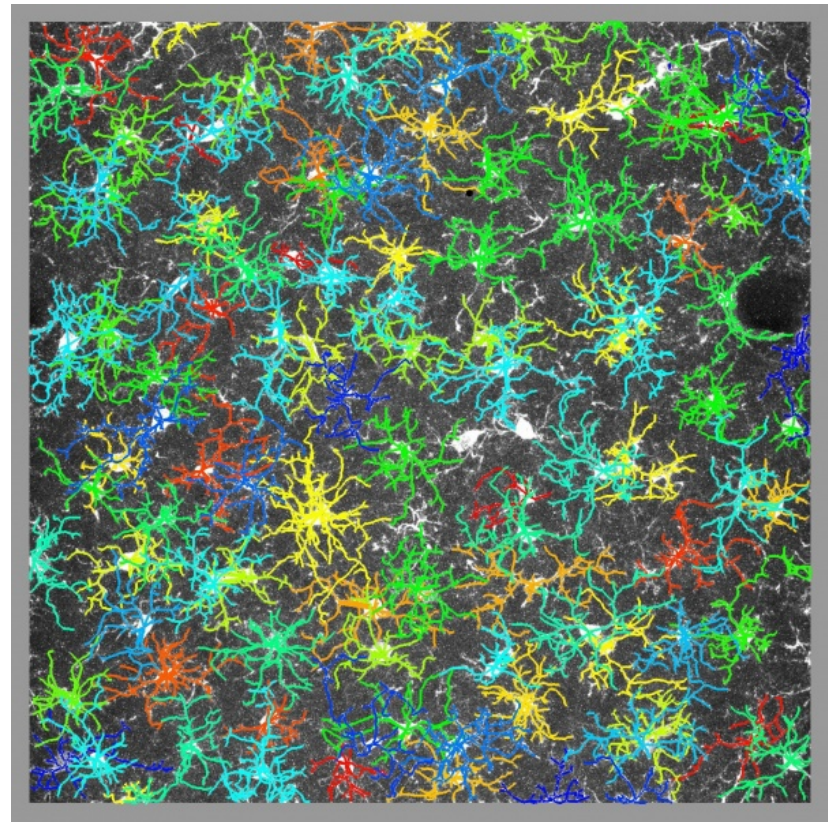


图 1.4 FARSIGHT 导出的重建结果

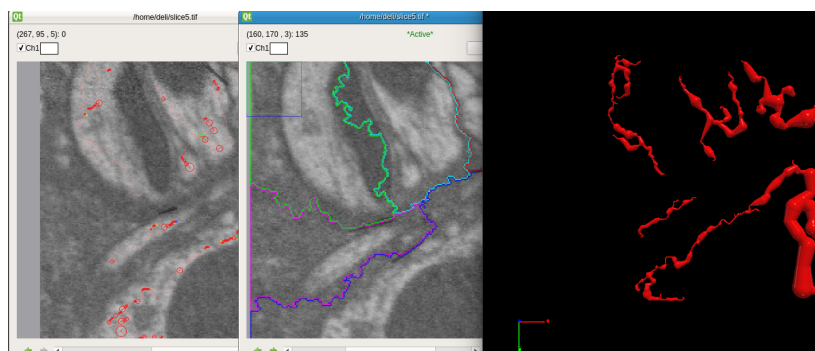


图 1.5 neuTube 软件运行界面

较高,但是如果遇到细小的非常见错误,FARSIGHT 无法识别,也不提供精细化的编辑手段,需要借助于其他软件完成。

1.2.2 neuTube

neuTube 是一种基于 SWC 文件格式的神经元重建软件,同时具备 2D 和 3D 的可视化以及直观地编辑、绘制功能,运行用户有效的根据荧光图像数据重建神经结构,并且编辑由其他软件生成的标准神经结构的文件^[2]。软件界面如图 1.5 所示。

虽然 neuTube 提供了 2D 和 3D 模式下精细编辑神经结构的功能,但是无法多人协同编辑,也无法保存编辑的历史记录,不利于多人协同工作。

1.3 论文结构

本文旨在设计并实现在线多用户的神经网络结构编辑分享平台,利用互联网便于数据共享与交流的特点解决一些现有神经元编辑软件的问题,使得神经科学家可以便捷地进行异地,多用户协同编辑神经网络结构,并能分享完成重建的结构脑图谱,探索神经元结构下的奥秘。由于项目涉及到数据可视化与后台服务器搭建,自然地将整个项目分成两部分,这里主要实现后台服务器的搭建,为前端可视化操作提供了有力的支持。

第一章讨论了交互式神经元重建系统的背景与意义以及神经元编辑软件的问题,第二章讨论了项目整体架构并简单介绍所用到的技术。第三章讨论技术实现的

细节以及如何根据性能测试报告进行性能优化,第四章描述了在性能优化之后的系统整体性能,第四章讨论了接下来的工作并分析了项目中存在的不足。

第 2 章 整体架构与技术选型

2.1 整体架构

项目整体结构如图 2.1 所示, 共分为神经信息数据库, 用户信息数据库以及网络服务器三部分组成。图中主要包含了两个方向的数据流, 一个是用户信息数据流, 另外一个神经信息数据流。

2.1.1 用户信息数据流

用户信息数据流主要负责三件事情, 一是验证用户身份, 二是获取用户资源列表, 三是检查用户行为是否有足够的权限。用户信息数据流可以抽象成如下几步:

1. 用户向网络服务器发送请求, 登录平台或注册新用户
2. 网络服务器向用户信息数据库验证用户身份或添加新的用户信息
3. 用户信息数据库返回验证结果给网络服务器
4. 网络服务器将结果反馈给用户

2.1.2 神经信息数据流

神经信息数据流负责维护三对关系, 用户和原始神经图片信息的关系, 用户和数字重建结果直接的关系以及数字重建结果和原始图片信息之间的关系。神经信息数据流可以抽象成如下几步:

1. 用户像网络服务器请求神经元信息
2. 网络服务器利用用户信息数据流中保存的用户信息返回对应用户资源列表
3. 用户信息数据库返回用户资源列表给网络服务器
4. 网络服务器根据用户资源列表向神经信息数据库查询对应神经信息
5. 神经信息数据库返回对应神经信息给网络服务器
6. 网络服务器返回用户所需的神经信息用于前端的可视化展示

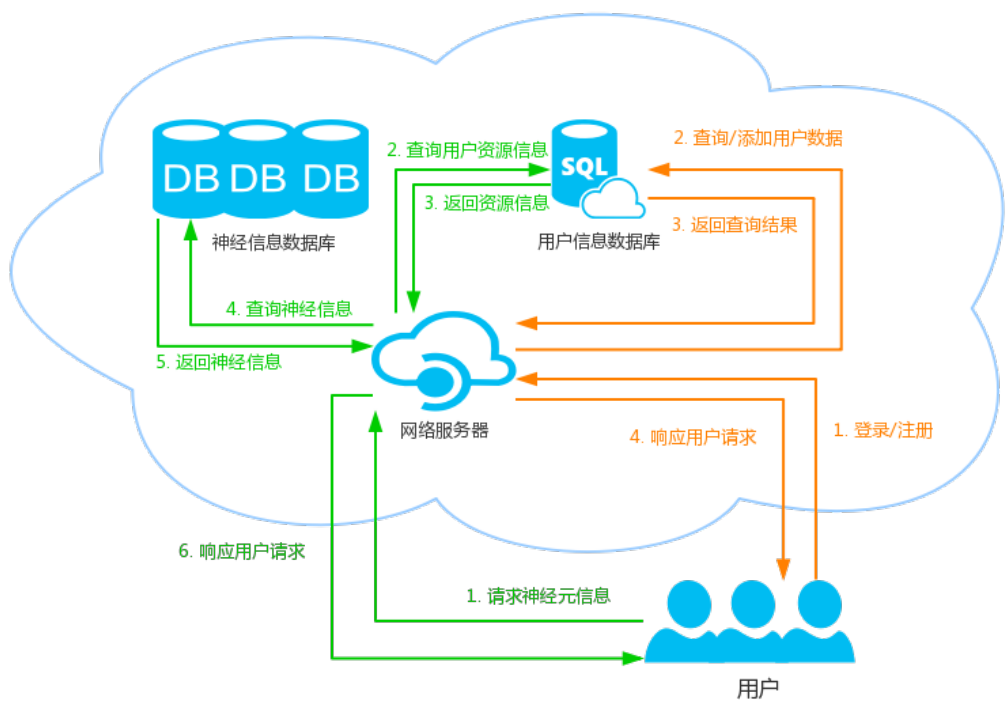


图 2.1 项目整体架构

2.2 技术选型

2.2.1 神经信息数据库

神经信息数据库包含两部分数据，一部分是原始大脑切片显微镜图像，另外一部分是初步完成数字重建的 SWC 文件。使用 DVID 作为数据库，储存这些信息。DVID 是一个分布式面向图像的数据服务，主要用于图像分析与可视化。DVID 有如下特点：

1. 便于扩展数据类型，允许用户根据数据特点加速访问速度，减少储存空间，提供方便的 API。这为储存数字重建结果提供了便利。
2. 为分布式数据储存提供了类似于 GIT 的版本控制系统，在此基础之上我们可以解决多用户同时编辑产生冲突的问题。
3. 方便连接其他 API 如 Google BrainMaps 和 OpenConnectome 等。
4. 支持多分辨率图像数据，使得用户可以在不同尺度下观察图像信息。

在 DVID 的基础上，构建出原始大脑切片显微镜图像与数字重建结果的储存仓库，将数据储存抽象成数据存储服务，使得可以专注于完成核心算法和逻辑。

2.2.2 用户信息数据库

用户信息数据库包含多用户管理以及用户资源管理。使用 PostgreSQL 数据库储存这部分信息。PostgreSQL 最初由加州大学伯克利分校计算机系开发完成。在支持大部分 SQL 标准之上，提供了许多诸如复杂查询，多版本并行控制，事物完整性等现代特性^[4]。由于 PostgreSQL 对标准 SQL 支持度较高，可以方便的和 DVID 联系起来，将用户信息和原始图像信息，数字重建结果对应起来。利用 PostgreSQL 支持的储存过程，事物以及多版本并行控制特性，我们可以方便的实现分布式，多用户实时编辑平台，并解决多用户同时编辑可能产生冲突的问题。

2.2.3 网络服务器

采用 Node.js 和 Express 完成网络应用开发。Node.js 是一个基于 Chrome V8 引起的 JavaScript 的运行环境。Node.js 使用了一个事件驱动、非阻塞式 I/O 的模型，使

其轻量又高效^[5]。Express 是一个基于 Node.js 平台的极简、灵活的 web 应用开发框架,提供丰富的 HTTP 快捷方式和任意排列组合的 Connect 中间件,帮助快速、简单的创建健壮、友好的 API。

第 3 章 平台搭建与性能调优

3.1 数据表设计

为了实现多用户管理以及用户资源管理，共设计实现了三张数据表，分别是用户信息表，原始图像数据表以及 swc 数据表。

- 1. 用户信息表用户信息表如表 3.1 所示，共有三个字段，分别储存了用户名，密码和用户保证账户安全的盐值。
- 2. 原始图像数据表原始图像数据表如表 3.2 所示，共有三个字段，分别储存了创建者，图像名和权限控制的用户角色。
- 3. swc 数据表原始图像数据表如表 3.3 所示，共有五个字段，分别储存了创建者，图像名，创建时间，swc 文件名以及用户评论。根据创建时间，可以建立同一图像下 swc 文件的拓扑顺序，为多用户同时编辑以及合并冲突分支提供了基础。

表 3.1 用户信息表

字段名	数据类型	备注
username	STRING	主键
password	STRING	
salt	UUID	用于保证用户账户安全

表 3.2 原始图像数据表

字段名	数据类型	备注
username	STRING	创建者, 外键, 用户信息表中的 username
image	STRING	
role	STRING	用于权限控制

表 3.3 原始图像数据表

字段名	数据类型	备注
username	STRING	创建者, 外键, 用户信息表中的 username
image	STRING	原始图像名, 外键, 原始图像数据表中的 image
createdAt	TIME	创建时间
swc	TEXT	swc 文件名
comments	STRING	备注

第 4 章 项目成果

4.1 第一节

填入正文内容

参考文献

- [1] Kerry M Brown, Germán Barrionuevo, Alison J Canty, Vincenzo De Paola, Judith A Hirsch, Gregory SXE Jefferis, Ju Lu, Marjolein Snippe, Izumi Sugihara, and Giorgio A Ascoli. The diadem data sets: representative light microscopy images of neuronal morphology to advance automation of digital reconstructions. *Neuroinformatics*, 9(2-3): 143–157, 2011.
- [2] Linqing Feng, Ting Zhao, and Jinhyun Kim. neutube 1.0: A new design for efficient neuron reconstruction software based on the swc format. *Eneuro*, 2(1), 2014.
- [3] Jonathan Luisi, Arunachalam Narayanaswamy, Zachary Galbreath, and Badrinath Roysam. The farsight trace editor: an open source tool for 3-d inspection and efficient pattern analysis aided editing of automated neuronal reconstructions. *Neuroinformatics*, 9(2-3):305–315, 2011.
- [4] Michael Stonebraker and Greg Kemnitz. The postgres next generation database management system. *Communications of the ACM*, 34(10):78–92, 1991.
- [5] Stefan Tilkov and Steve Vinoski. Node.js: Using javascript to build high-performance network programs. *IEEE Internet Computing*, 14(6):80–83, 2010.
- [6] Lav R Varshney, Beth L Chen, Eric Paniagua, David H Hall, and Dmitri B Chklovskii. Structural properties of the caenorhabditis elegans neuronal network. *PLoS Comput Biol*, 7(2):e1001066, 2011.
- [7] John G White, Eileen Southgate, J Nichol Thomson, and Sydney Brenner. The structure of the nervous system of the nematode caenorhabditis elegans. *Philos Trans R Soc Lond B Biol Sci*, 314(1165):1–340, 1986.

致 谢

附 录

一、题目：_____

二、指导教师对毕业论文(设计)的进度安排及任务要求：

起讫日期 200 年 月 日至 200 年 月 日

指导教师(签名)_____ 职称 _____

三、系或研究所审核意见：

负责人(签名)_____

年 月 日

毕 业 论 文 (设计) 考 核

一、指导教师对毕业论文(设计)的评语:

指导教师(签名) _____

年 月 日

二、答辩小组对毕业论文(设计)的答辩评语及总评成绩:

成绩比例	文献综述 占 (10%)	开题报告 占 (20%)	外文翻译 占 (10%)	毕业论文 (设计) 质量及答辩 占 (60%)	总评成绩
分值					

答辩小组负责人(签名) _____

年 月 日