**中图分类号： 单位代号：10280**

**密 级：公开 学 号：17721758**

**硕 士 学 位 论 文**

## SHANGHAI UNIVERSITY

**MASTER’S DISSERTATION**

|  |  |
| --- | --- |
| **题**  **目** | **基于深度学习的三维点云手部姿态估计** |

**作 者 阎冲**

**学科专业 计算机应用技术**

**导 师 张景峤**

**完成日期**

姓 名： 阎 冲 学号：17721758

论文题目：基于深度学习的三维点云手部姿态估计

上海大学

本论文经答辩委员会全体委员审查,确认符合上海大学硕/博士学位论文质量要求。

答辩委员会签名：

主任：

委员：

导 师：

答辩日期：

姓 名： 阎 冲 学号：17721758

论文题目：基于深度学习的三维点云手部姿态估计

**原 创 性 声 明**

本人声明：所呈交的论文是本人在导师指导下进行的研究工作。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已发表或撰写过的研究成果。参与同一工作的其他同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签 名： 日 期：

**本论文使用授权说明**

本人完全了解上海大学有关保留、使用学位论文的规定，即：学校有权保留论文及送交论文复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容。

（**保密的论文在解密后应遵守此规定**）

签 名： 导师签名： 日期：

上海大学工学硕士学位论文

**基于深度学习的三维点云手部姿态估计**

姓 名：阎冲

导 师：张景峤

学科专业：计算机应用技术

上海大学计算机工程与科学学院

2020年6月

A Dissertation Submitted to Shanghai University for the Degree of Master of Philosophy

in Engineering

**The Research of a New NC Code**

MA Candidate：Lixin Lu

Supervisor：Jingqiao Zhang

Major：Computer Applied Technology

**School of Computer Engineering and Science,**

**Shanghai University**

**×××, 2005**

摘 要

… …

（注：简要介绍本论文的主要内容，主要为本人所完成的工作和创新点）

**关键词：**（注：不宜多，最好不超过5个）

ABSTRACT

**Keywords:**

目 录

[SHANGHAI UNIVERSITY 1](#_Toc12661)

[摘 要 5](#_Toc11870)

[ABSTRACT 6](#_Toc20552)

[目 录 7](#_Toc14474)

[第一章 绪论 1](#_Toc3571)

[1.1 课题研究背景和意义 1](#_Toc15004)

[1.2 国内外研究概况 2](#_Toc26537)

[1.2.1 生成方法 3](#_Toc23128)

[1.2.2 判别方法 3](#_Toc7155)

[1.2.3 混合方法 3](#_Toc13468)

[1. 2. 4 基于深度学习的手部姿态估计 4](#_Toc4838)

[1.3 本文主要研究内容 5](#_Toc32036)

[第二章 相关 6](#_Toc21941)

[2.1 数据集 6](#_Toc20878)

[2.1.1 NYU数据集 6](#_Toc24463)

[2.1.2 ICVL数据集 6](#_Toc26138)

[2.1.3 MSRA数据集 6](#_Toc24896)

[2.2 评估指标 6](#_Toc8743)

[第三章 基于KPconv的深度学习框架 7](#_Toc27738)

[3.1 神经元 7](#_Toc15162)

[3.1.1 基本介绍 7](#_Toc2461)

[3.1.2 数据来源及说明 7](#_Toc25375)

[3.2 预处理 8](#_Toc26503)

[3.2.1 神经突片段 8](#_Toc19123)

[3.2.2 神经突中心曲线 9](#_Toc9847)

[3.3 表面网格构造 11](#_Toc16390)

[3.3.1 片段表面网格生成 11](#_Toc30395)

[3.3.2 片段表面网格拼接 13](#_Toc27298)

[3.3.3 胞体表面网格生成 13](#_Toc8264)

[3.4 实验结果和分析 14](#_Toc32506)

[3.4.1 实验环境和方法 14](#_Toc2279)

[3.4.2 实验结果和分析 15](#_Toc19015)

[3.5 小结 15](#_Toc13599)

[第四章 手部姿态动态跟踪系统 16](#_Toc16994)

[4.1 粗糙网格局部优化 16](#_Toc24871)

[4.2 网格细分 16](#_Toc24923)

[4.3 数据结构 16](#_Toc31393)

[4.3.1 几何形体表示 16](#_Toc4831)

[4.3.2 半边数据结构 17](#_Toc18667)

[第五章 结论和展望 18](#_Toc27044)

[5.1 流程的制定 18](#_Toc16717)

[5.2 CNCS的应用 18](#_Toc17791)

[5.3 CVS的应用 18](#_Toc11604)

[5.4 结论 18](#_Toc5832)

[第六章 结论与展望 19](#_Toc15621)

[6.1结论 19](#_Toc4660)

[6.2展望 19](#_Toc11233)

[作者在攻读X士学位期间所作的项目 22](#_Toc95)

[致 谢 23](#_Toc5834)

第一章 绪论

## 课题研究背景和意义

在当今社会，人机交互技术在日常生活中的作用越来越突出，如何更高效地实现人和机器设备之间的交互操作，一直是人们不断努力提升的重要领域。其中最具代表意义的就是鼠标和键盘的发明，伴随着计算机诞生之日起到现在，成为了人们办公必不可少的输入设备。在前几年的手机设备上，大多数仍然以按键作为输入，当触摸屏技术出现之后，物理按键的这种操作方式便很快被人们摒弃，从市场的主流到消失在大众的视野里，也只用了短短的几年时间。可以看出，便利快捷的交互方式，始终是人们所追捧的。

手作为人体最灵活的部位，在承担了生活中绝大多数劳动工作的同时，也在人与世界的交互过程中起着重要的作用。所以，绝大多数人与机器之间的交互都是通过手来完成的。但是这些任务往往都是利用手去进行实际操作，并未用到手本身更加便利、自然的交互方式：三维手势技术。三维手势技术往往不需要进行实际的接触操作，仅通过姿态的变化就可实现丰富的手势，所以三维手势可以具有非常广泛的应用。当下最火的应用当属虚拟现实技术，业界将 2016 年称为虚拟现实技术（VR）元年， CES 电子消费展上 Oculus、三星、HTC 等各大厂商纷纷推 出了自己的 VR 产品，诸如 Oculus 的 RIFT、HTC 的 VIVE等 VR 产品，三维手势将是其最理想的输入方式之一。在未来十年，随着 VR 技术和设备的井喷式发展，三维手势技术必然能够拥有一席之地。除了 VR领域之外，家用机器人产业成为三维手势的另一个主要应用领域。传统的机器人功能往往比较简单，难以完成复杂的作业任务。随着近几年计算机视觉技术、导航技术、大数据和深度学习领域所取得的突破，面向家用的机器人领域也将迎来新的机遇。目前机器人技术需要一种或多种交互技术，来提升广大用户的交互体验。如果让机器人能够像人一样，通过识别人们的手势与人进行交互，这才是更加理想的方式。国外在20世纪90年代就开展了相关研究，其中[1][2]等文提出通过机器人对手势方向进行检测，实现机器人的引路功能，而文章[3]则实现了人与机器人更加自然的交互方式，通过机器人识别用户的手势来理解用户想表达的含义和心情，使得交互方式更加人性化，而在[4]中进一步将手势和头的面朝方向信息相结合，实现更加拟人化的交互理解。

除此之外，手语识别、互动娱乐、桌面交互等领域也是三维手势的重要应用场景，三维手势技术的发展对这些领域起着巨大促进作用。

三维手势研究涉及从手部信息的捕获到手势的识别的各个方面，主要技术分为手检测、手跟踪和手势识别三部分[5]。其中，手检测是指从获取的数据中如何实时获得手部区域信息，为手势识别打下基础。手跟踪则是实时获取当前手势的姿态信息，这一步骤也可称为手部姿态估计，其主要目的是利用手部数据去估计当前手势动作下的手部姿态。手势识别主要侧重于对用户手势含义的识别，进一步可以分为静态手势、动态手势以及轨迹手势等类型。

手部姿态建模技术同样是三维手势领域的重要技术，其实现了手部姿态的实时估计，它的结果决定了手势识别以及相关应用的实际效果。在相关设备方面，手部姿态估计最开始使用的是接触式设备，即用户通过佩戴数据手套[6]，以检测不同手势下的手部姿态参数。这种设备昂贵，而像佩戴颜色手套[7]等这类方法极大降低了用户的舒适便利性，因此主流方法从使用接触式设备转移到了计算机视觉。按照综述文章[8][9]的方法分类，最初的研究方向始于基于表观法的静态手势识别领域，即利用拍摄的图像分割手部区域之后，检测轮廓曲率特征[10]、指尖点位置[11]等信息，对其结果进行模板匹配或设定约束来实现手势识别；也有文章[12]对手部区域提取 Haar-like 特征并利用 Adaboost 分类器进行手势识别。这样的静态手势识别多采用彩色图像，能够估计出姿态比较特殊的手势模板参数，但是在实际场合下往往背景比较复杂，同时手部的自由度较高、外观形变大，这些因素使得基于静态手势识别的方法并没有对交互领域起到推动作用。

最近几年，深度摄像机的普及和流行颠覆了很多领域的传统方法，同时也使得一些新领域成为研究热点。当下比较火热的研究领域包括姿态估计、物体识别以及地图导航等。以姿态估计作为代表，手部姿态估计技术已经得到极大的提升，微软推出的 Kinect 传感器已经在多个领域代替彩色相机展开应用。

Kinect 传感器相比传统的彩色摄像机，能够以廉价的方式获得拍摄物体相对于相机的深度信息，即深度图像。根据获得被拍摄物体的深度图像，不仅可以更方便的分割背景和被拍摄物体，而且相比传统的 2D 图像数据更能够表达物体在形变过程中表面所发生的实际变化。因此，基于深度数据的手部姿态建模技术，逐渐占据了主流。Kinect除了可以获取深度图像数据之外还可以获取点云数据，以点云数据形式作为手部姿态估计的处理数据，可以更好地保留手的3D信息。

## 国内外研究概况

手部姿态估计技术可分为三种，生成方法、判别方法和混合方法。其中，生成方法被称为基于模型的方法，判别方法被称为基于外观的方法，而混合方法则由这两种方法融合而成。这些方法都是读入深度图像输出对应的手势模型。

生成方法的核心思路是：先创建大量的手势，之后于当前深度图像中选择最匹配的手势，基于输入的深度图像与手模型深度图像的相似性，对目标函数进行优化，进而找到最接近的手模型。该类方法的优点是可以输出高精度的手势，缺点则是计算量大，估计的错误容易出现累积。

判别方法的核心思路是：学习从深度图像到手势的映射。优点是由于每帧都会进行初始化过程，因此错误不会传播到下一帧，缺点是深度图像的低分辨率、自我遮挡、快速移动会产生大量错误。

相比于生成方法的错误累积缺点，判别方法每帧都会初始化手势的这一特点可以让整个估计过程更容易从追踪失败中恢复正常。在呈现效果方面，由于从大量数据中学习而得的判别模型只能估计出一个粗糙的手势，而生成方法则可以通过使用高质量的手部几何模板来提高呈现效果。

* + 1. 生成方法

生成方法通过图像的时间序列来匹配模板[13][14][15]，如果给定所跟踪手的精确模板，这些方法可以追踪更精细的手部动作。Oikonomidis等人[16][17]则使用了高质量的3D手势模板来进行关节运动中的手势跟踪和精准3D姿势估计[6]。

Oikonomidis等人[13]提出了一种基于粒子群优化（PSO）的生成方法，可以使用深度传感器进行完整的DoF手部跟踪（15Hz）。该方法使用初始姿势模型跟踪手的动作，但无法从跟踪丢失中恢复。钱等[19]通过增加一个预处理的 PSO步骤和重新初始化程序来扩展[13]。 Melax等人[14]和钱等人[19]的简单几何初始化试探法可以克服从上一帧优化中造成的跟踪丢失问题。 Melax等人[14]使用由物理处理器驱动的生成方法进行3D姿态估计。这些方法所处理的场景距离范围小且使用的是简单的多面体模型，而Fitzgibbon等[20]的工作范围可扩展到数米范围，使用了更完整的3D网格模型。

* + 1. 判别方法

与生成方法不同的是，判别方法是通过从大量注释样本数据集中学习如何从每个图像中提取特征来估计手势[21][22][23]。判别方法避免了生成方法易出现的每帧估计偏移问题，但最终效果没有生成方法的精确，而且判别法的关节估计容易违反运动学约束。判别方法读取图像数据（例如，提取图像特征并使用分类或回归技术），建立总图像数据到一组预定义手势姿势的映射。因为判别方法通常不需要时间信息，所以它的重新初始化阶段增强了姿态估计过程的鲁棒性[24]。

手部姿态估计的判别方法[21][22]往往需要大量的训练样本进行训练。训练中的学习方法有增强树分类器[25]、随机决策森林[21]和随机森林回归[23][26]方法等。Keskin等[21]使用多层随机森林来预测手部姿态匹配简单骨架。该系统在消费者级别的CPU硬件上以30Hz的频率运行，但在遮挡下易发生错误。唐等[22][27]扩展了这项工作，可以在25Hz频率下跟踪更加复杂的手部姿态，并在手自我遮挡方面比Keskin等[21]更具鲁棒性。但由于两种方法都没有明确的模型匹配步骤，导致结果显示的手部姿态有时并不协调（例如，关节反转或手指长度不协调）。徐等[28]估计出手的整体方向和位置信息，并通过最小化重建误差来选择匹配的手势，追踪有21个自由度的手部姿态。但该方法的运行频率仅为12Hz，估计过程中容易引起抖动。汤普森等[29]的手部跟踪效果好，其使用深度神经网络来预测手部特征位置、使用IK方法判断骨架。该方法实时效果显著，但仅能处理近距离跟踪。Wang等[30][31]使用大型手势数据库的判别式近邻查找方法，并用IK方法改善手势效果。最近邻方法高度依赖于手势数据库，很难处理数据库之外的手势。

* + 1. 混合方法

将上述两种方法进行结合会很大程度的优化他们的问题。混合方法[19][22][32]首先利用判别框架从每帧估计出粗略的手势，然后使用生成方法将其匹配对齐到所使用的手部模板。Sridhar等人[49]和Fleishman等人[50]通过由像素森林模型得到的标记加强每帧数据与手模板的对应关系。而Sridhar等人[51]则将基于判别方法的手势检索与基于生成方法的跟踪过程结合起来进行手部姿态估计。在训练数据方面，Rogez等人[52]使用计算机图形模型来生成训练数据，将其用于学习特定姿势的判别模型训练。

* + 1. 基于深度学习的手部姿态估计

用深度学习的方法进行手部姿态估计，大致可以分为两个过程：首先通过CNN等方法提取图像特征，进而获得相应关键点的热力图，最后通过最大化后验概率进行优化。但在研究过程中，学者们针对相应的问题又提出了一些新颖的方法，例如基于多摄像头和多视图的创新方法。

Jonathan等人[32]提出一种使用CNN提取特征，并为关节位置生成热图，最后从特征和热图中使用IK（inverse kinematies）算法推断出手的姿势。该方法中，使用热图（heatmap）只能预测出关节点的2D位置，而第三维坐标是通过深度图像计算得到的。Markus等人[33]提出了一种可以直接从给定深度图像中预测出手的3D关节位置的卷积神经网络架构。该方法设计一个预估网络输出手的关节点位置，再把每一个关节点再输入到另一网络中进行3D位置的优化。两个网络的结合，使得系统效果在时间和精度上都有了很大的提升。Ge等人[34]提出了一种多视图CNNs的手部姿态估计方法。该方法把深度图像的三视图同时输入到卷积神经网络中，产生前视图、侧视图和俯视图对应的热图，再把三个热图融合起来得到最后的3D手部姿态估计。该方法利用三个视图的热图缓解了评估的模糊问题，省去了复杂的模型矫正过程。

除了二维卷积神经网络之外，一些三维卷积神经网络（3D-CNN）模型被提出应用于手部姿态估计。Molchanov等人[35]提出一个3D-CNN模型：通过深度数据和强度数据对汽车驾驶手势进行初步识别，再将多个空间尺度的信息组合起来进行最终估计。Molchanov等人[36]使用循环机制对3D-CNN模型进行拓展，对动态手部姿态检测和分类。其网络模型包含一个提取时空特征的3D-CNN模型和一个针对全局时序建模的循环层。Li等人[37]对Tran等人[38]提出的3D-CNN模型进行改进，使用深度图像与RGB视频进行大规模的手部姿态估计。同样地，Camgoz等人[39]也以Tran模型为基础构建用于大规模手部姿态估计的端到端（end-to-end）3D-CNN模型。Miao等人[40]将3D-CNN与ResNet结合提出ResC3D模型，从时序RGB图像、光流数据、深度数据提取手部特征，通过SVM进行手势分类。该方法在Chaleam、LAP、IsoGD数据集上取得67.71%的识别率。马正文等人[41]则是使用稠密连接的3D-CNN进行手部姿态估计。

应用于手部姿态估计领域的深度学习除了可以处理深度图像之外，还能够对点云数据进行处理。

2016年，Qi等人[42]首先提出了直接处理3D点云的深度学习框架PointNet，并在2017年又提出了改进版PointNet++[43]。该框架主要处理的是点云的分类分割问题。而在pointnet之前，深度学习方法对点云数据的处理通常是先将点云做预处理成其他数据形式之后再训练：常见方法是将点云转换成多个视角的2D图片或将点云表示为空间中自带依赖关系的体素。随后，Li等人[44]引入3D-CNN的思想处理点云。他们从输入的点云中学习X变换矩阵进行卷积操作。其中，该方法提出的X矩阵在理想条件下综合考虑了点云的形状的同时不依赖点云中点的顺序，可以在训练过程中较为完整的保留点云的三维信息。从此开始，手部姿态估计也有了基于PointNet处理三维点云的深度学习框架[45][46]。Liu等人[45]基于PointNet提出一种可以捕捉复杂手部结构的回归网络，并引入指尖优化网络细化指尖的位置。除此之外，Liu等人[46]还提出了另一种基于PoinNet的方法：用PointNet预处理点云，对点云的热图和单位矢量场处理得到最终的手部姿态。Chen等人[47]提出了深层语义回归网络SHPR-Net，将先验语义和后期融合集成在一起对手部姿态估进行回归。Li等人[48]方法将点云数据优化为1024个点，使用基于投票的方法训练模型，完成对手各部分的分割。

## 本文主要研究内容

本文主要对手部姿态估计的问题进行了研究。在手部姿态估计领域，国内外的研究人员对使用点云作为输入的深度学习框架研究不多。鉴于点云这一数据格式对三维空间信息的表征更加具体，有必要对点云学习方法进行更加深入的研究。本文提出了点云学习框架应用于手部姿态估计领域的方法。本方法首先使用Deepsets处理点云所存在的置换不变问题。接着对点云进行局部区域划分，并提取点云的局部特征、全局特征。最后融合点云的局部特征和全局特征，通过全连接层输出手部的关键点信息。除此之外，本文还提出了手部姿态动态跟踪系统。该系统使用Kinect采集实时三维空间数据，输出跟随手部移动的人手模型。

本文章节安排如下：

第一章首先介绍手部姿态估计研究的背景及意义，阐述了该技术的国内外研究现状。

第二章简要概述手部姿态估计任务中常用的2.三类数据集和实验结果3.常用的评估指标、1.基础性的理论（pointnet基本框架，第三章要用到的基础理论）。

第三章介绍本文提出的基于KPConv的手部姿态估计深度学习框架。

（丰富的话-->实验单章）or更进一步模型or平行方法

第四章介绍手部姿态动态跟踪系统。该系统使用kinect传感器作为数据的来源，将从Kinect得到的深度图像分割出手的位置并转换为点云，输入到上一章的深度学习框架中，利用得到的手部关键点信息驱动手的三维模型，对三维空间中人手进行跟踪。（具体的对比、效果、架构、仅为补充）

第五章总结本文提出方法的优势和不足，对基于点云的手部姿态估计技术进行了展望，提出今后的研究内容。

（均要承上启下，本章小结启下引出下一章章节开头承上，节与节之间。故事线贯穿联通）

第二章 相关

## 数据集

* + 1. NYU数据集
    2. ICVL数据集
    3. MSRA数据集

## 评估指标

* + 1. PJME

per-joint mean error：第一个指标是关键点平均误差，指的就是在所有测试集中，实验结果和对应测试集关键点的平均误差距离。

* + 1. PGF

proportion of good frames：第二个指标是在测试集中合格项所占的比例（不分小结，为了叙述方便，可用自己缩写）

第三章 基于KPconv的深度学习框架

## 神经元

* + 1. 基本介绍

神经元是大脑结构和功能的基本组成单位。通过细胞染色技术，可以反映出神经元主要是由胞体和神经突起组成。神经突起包括轴突和树突两种。轴突是神经细胞所特有的结构，每个神经元只发出一条，其长短不一，表面光滑，分支少，且粗细均匀，一般会以直角发出少数侧枝，称为轴突侧支。轴突的作用是传出胞体所发生出的神经冲动信息。树突是从胞体延伸出来的细管状分支，类似于一棵树上长出的分枝，一个神经元的树突统称为树突树。树突上也常有分支，称为树突侧支。树突的大小、形态和数量与神经元的发育和功能相关。单个神经元的基本结构如图3.1所示。神经元形态的追踪重建的主要目的就是在图像堆栈中提取出上述结构，并数字化为三维采样点数据，这些数据为我们的表面构建算法提供了基本的信息。

* + 1. 数据来源及说明

本文实验所用的数据可以在迄今为止最大的、可公开访问的3D神经元重建及相关元数据的数据库NeuroMorpho.Org中得到。该数据库中集中收集了很多来自全球多个实验室，并经过同行认可的数字化重建的神经元信息。SWC格式是该数据库用来存储数字化后的神经元结果的方式之一。表格1中给出了使用这种格式描述神经元结构的示例。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| # | id | t | x | y | z | r | p |
|  | 1 | 1 | 0.0 | 0.0 | 0.0 | 8.4645 | -1 |
|  | 2 | 1 | -4.11 | 7.4 | 0.0 | 8.4645 | 1 |
|  | 3 | 1 | 4.11 | -7.39 | 0.0 | 8.4645 | 1 |
|  | 4 | 3 | -1.01 | 2.2 | 0.0 | 6.2016 | 1 |
|  | 5 | 3 | -1.71 | 3.9 | 0.0 | 2.7669 | 4 |
|  | 6 | 3 | -2.74 | 5.88 | 0.0 | 2.3036 | 5 |
|  | 7 | 3 | -3.38 | 8.2 | 0.0 | 21626 | 6 |
| …… | | | | | | | |
|  | 80 | 3 | -2.19 | 7.96 | 0.0 | 1.6723 | 6 |
|  | 81 | 3 | -1.93 | 9.49 | 0.0 | 1.4510 | 80 |
| …… | | | | | | | |
|  | 205 | 2 | 2.22 | -4.31 | 0.0 | 2.4826 | 1 |
|  | 206 | 2 | 3.49 | -5.64 | 0.0 | 1.9950 | 205 |
| …… | | | | | | | |

表格1中的数据属于老鼠脑干细胞，具有胞体、轴突以及树突结构域，在NeuroMorpho.Org的标识编号为NMO\_45927。表中的每一行都表示一个神经元形态采样点，附带有七个说明结构信息的数据项，分别是编号*id*，类型*t*，坐标*x*、*y*、*z*，半径*r*和父采样点的编号*p*。

对于一个形态学采样点，类型指的是该采样点所属的结构域，即胞体、轴突或者树突；空间坐标反映了该采样点在三维空间中的位置，半径则反映的是该采样点所属神经突厚度的一半。这些采样点是离散存储的，但采样点之间存在连接关系，这种关系根据父采样点编号确定，即空间中相连的两个邻近采样点之间具有父子关系。根据这种父子关系，可以将这些采样点连接在一起，构成描述神经元形态的中心骨架折线状结构，即形态骨架，如图3.2所示。

## 预处理

首先，我们将所有的形态采样点信息从SWC文件中读取并存储起来。然后，根据这些从文件中得到的采样点生成一些表示不同神经突的片段。将数据分成片段进行处理能够简化网格的生成过程。

* + 1. 神经突片段

上文提到，神经元由胞体和从胞体辐射出的神经突起组成，这些神经突起可以看作是一个个分支。从文件中读取到的内容就是这些分支的离散表示。在本小节，将对所有的数据进行分段处理，以区分哪些采样点组成的序列表示一个突起分支，即神经突片段。

除了表示胞体的采样点之外，我们将其余采样点分为三类：普通点，分叉点和末端点，如图3.3所示。在对数据进行分段处理时，需要结合采样点的类别，并根据采样点之间的父子关系决定哪些点组成一个片段序列。具体的规则是：如果当前采样点是普通点，则该点与其子采样点属于同一个序列；如果当前采样点是分叉点，则需要根据该点前后相关采样点的空间分布情况决定哪一个子采样点与它属于同一个序列。当连接分叉点的父采样点、分叉点和分叉点的子采样点的折线段形成的夹角接近平角时，该局部区域表面网格的构造会更加容易和方便。因而，可以认为满足此条件的子采样点与当前分叉点属于同一个序列，其它的子采样点则引导出新的序列。至于末端点，它是一个片段序列中的最后一个采样点。

经过上述分段处理，除胞体外的所有采样点被分配到不同的片段序列当中。为后文叙述方便，我们根据序列中第一个采样点的父采样点的类型将这些片段人为地分为两类：S型片段和B型片段，分别记作S-seg和B-seg。所有S型片段的第一个采样点的父采样点是胞体，表示这条片段是实际上连接在胞体上的；所有B型片段的第一个采样点的父采样点是分叉点，表示这条片段实际上是从其它神经突起片段上延伸出的分支。

* + 1. 神经突中心曲线

经过分段处理后，单个神经元的神经突起可以看作是多个片段的集合，每个片段是一系列离散采样点的集合。由于神经元的形态骨架是由多段折线段构成，可视化效果不如连续曲线。此外，在后文中的表面网格生成算法需要计算中心骨架线上某处的切向量。因此，有必要使用连续曲线代替折线段。

对于每一个片段，构造一条穿过片段所有采样点的曲线。在数学上，将通过一组给定的离散点的曲线称为插值样条曲线。本文中使用Catmull-Rom插值方法为每个片段计算相应的样条曲线，这是因为该算法无需求解方程系统就可以直接插值于控制顶点，其主要特点是插值生成的曲线刚好通过所有控制点，而本文中组成一个片段的采样点可以被看作是控制点。

Catmull-Rom样条曲线拟合方法至少需要四个控制点，公式如下：

（3.1）

其中，，该区间内的每个值都对应所求曲线中的某点。实际上，如果只有四个控制点，曲线只会经过中间两个控制点，如图3.3所示：

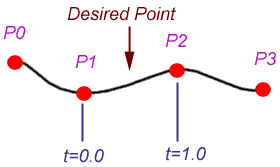


图3.3 Catmull-Rom样条曲线

在图3.3中，根据公式3.1可以得到连接和的样条曲线段。按照这种计算方式，需要引入两个新的控制点才能得到连接、、和的样条曲线段，这是使用该算法需要满足的必要条件。因此，对于一条由若干个采样点组成的神经元片段，在计算相应的插值曲线时也需要添加额外的采样点以满足计算条件。

在3.2.1小节中，我们根据每个神经突片段第一个点的父采样点的类型去区分该片段是从胞体还是分叉点处延伸出的分支，因而，父采样点可以自然地作为额外添加的点之一。另一个需要引入的点则直接使用每个片段的末端点进行表示即可。

通过Catmull-Rom插值算法得到每个片段相应的插值曲线之后，形态骨架从折线表示变为曲线表示，可视化效果更加自然。从而，骨架曲线上任一点处的切向量可以通过计算插值曲线在该点处的一阶导数求得。



图3.1 系统硬件平台

## 表面网格构造

实验所构造的表面网格，是以形态骨架为中心轴线的封闭表面。首先，为各个神经突片段以及胞体生成表面网格；然后，对已生成的表面网格进行拼接，拼接过程包括不同神经突片段的表面网格拼接以及胞体和神经突片段表面网格的拼接。

* + 1. 片段表面网格生成

每个神经突片段的表面网格，类似于一段沿着神经突中心曲线的柱体管道。上文中组成神经突片段的采样点离散分布在中心曲线上，因此，每个采样点处都有一个横截面，垂直于中心曲线在该点处的切线向量。我们需要计算这些横截面并进行连接，从而得到单个神经突片段的表面网格。在本实验中，我们使用正方形来近似表示这个横截面。

假设一个神经突片段由个离散采样点组成，记为。我们需要在处构造一个正方形，本文称其为采样点处的轮廓。轮廓是以采样点为中心，的厚度为半径的圆的外接正方形，其所在的平面是插值曲线在采样点处的法平面，如图5a所示。

在一个片段中，除了第一个采样点之外，其它每个采样点处的轮廓顶点都是基于其前一个采样点，即父采样点的轮廓顶点计算得到。如图5b所示，点是采样点处的轮廓的第一个顶点，则点处的轮廓上的对应顶点可如下计算：

（3.2）

式中，***T***，***P***，***N***分别代表不同的向量操作。***T***表示将向量沿着方向平移，平移距离为；***P***表示向量投影操作，将平移后的向量投影到点处的法平面上；***N***表示向量标准化操作。将标准化后的投影向量乘以缩放因子*s*即可得到向量。

采样点处的其它顶点可以根据罗德里格斯旋转公式得到。罗德里格斯旋转公式是计算三维空间中，一个向量绕着旋转轴旋转给定角度以后得到新向量的计算公式。该公式的数学化描述如下：

（3.3）

式中，v表示待旋转的三维向量，k是单位化的旋转轴向量，是旋转的角度，是旋转后的向量。在本方法中，将作为给定的向量，插值曲线在采样点的切向量作为旋转轴。旋转后的向量从指向。变换公式中的，可以得到不同位置的轮廓顶点。通过这种方式，分别为片段上除第一个采样点之外的其他采样点计算对应的轮廓。至于采样点处的轮廓，则需要按照轮廓顶点单独指定。

得到神经突片段的每一个采样点处的轮廓后，依次连接相邻轮廓从而生成片段的表面网格。虽然这些轮廓大小不一，相互之间也不平行，但控制公式3.3中的旋转角度相同时，相邻轮廓的顶点是一一对应的。这就使得连接规则变得简单，只要在生成网格时，对应连接相邻轮廓的顶点即可。

* + 1. 片段表面网格拼接

算法在3.2.1中对神经元数据进行处理得到神经突片段并在3.3.1中分别生成了这些片段的表面网格。不同神经突片段的表面网格可能会在分叉点附近产生相交，出现交叉重叠现象，如图3.4所示。如果直接进行拼接，则拼接后的网格在分叉点附近可能会产生奇怪的外观表现，例如，一些用于连接的线条穿插在网格面片之间。因此，在不同片段的表面网格进行拼接时，需要消除这些相交以避免网格重叠。

以下以两个神经突片段S-seg和B-seg之间的连接为例进行说明。根据上文描述的片段区分规则，此处实际上是将B-seg的表面网格拼接到S-seg上。首先使用包围盒进行相交性检测。在分叉点附近生成一些包围盒将相关表面网格包裹在里面。如图3.4所示，是S-seg上的包围盒，是B-seg上的包围盒。如果和任一有相交，则删除内包裹的网格面片，删除完成后的结果如图3.4所示。

显然，B-seg的第一个采样点的父采样点是S-seg中的采样点，后者本质上是S-seg中的一个分叉点。两个神经突片段之间的连接实质上是在S-seg的表面网格上找到一个合适的面，该面面向B-seg的第一个采样点处的轮廓。我们采用“射线法”找到这个合适的面。

以B-seg中的第一个采样点为起点发出一条射线，指向该点的父采样点，即S-seg中的分叉点。在射线前进的路径上，至少存在一个分叉点附近的网格面片和射线有交点。分别计算射线起点到这些交点的空间距离，并标记出和射线起点距离最近的交点所在的面片。然后，连接标记面片和轮廓的顶点形成新的网格面片，填充两个片段之间的空隙，如图3.4所示。（射线和空间有界平面的交点）

* + 1. 胞体表面网格生成

在生物学上，胞体表现为一个神经元的活动中心，可以接受其它神经元传来的冲动，也负责将发出传递到其它神经元的冲动。然而，大部分神经元追踪重建算法在执行的过程中，主要是跟踪神经突起的结构，将胞体作为区分不同神经元的一个重要特征。因此，在数据中，胞体仅仅表现为一个形态学采样点，是各个S-seg的交汇点。在本文中，使用一个四边形网格表示的类球体来表示这个中心。这个类球体可以通过对标准正方体进行一次Catmull-Clark细分得到。Catmull-Clark算法是一种针对四边形网格的细分法则，在执行过程中，该算法为每个面片生成一个新的顶点，每条边计算生成一个新的顶点，同时更新每个原始顶点的坐标位置。细分前初始正方体的边长是根据连接在胞体上的所有S-seg的第一个采样点的半径属性综合决定，防止其过大或者过小。

一旦生成了神经突片段和胞体的表面网格，很自然地需要将它们连接在一起。由于胞体网格面片是四边形，因此只需要找到合适的面，就可以直接和S-seg的第一个采样点处的轮廓进行拼接。这个过程类似于不同神经突片段之间的连接。

首先进行相交性检测。由于胞体是用类球体表示的，且在处理时，其中心位于坐标原点，因此检测方法可以简化为判断构成片段采样点轮廓的顶点是否位于该类球体内部。如果是，则删除对应的采样点轮廓以及基于该轮廓形成的表面网格面片。然后，使用“射线法”找到欲进行拼接的胞体网格面片并标记。连接标记面片和片段第一个采样点的轮廓顶点形成新的网格面片，填充片段和胞体之间的空隙。在胞体处的网格拼接过程中，使用“射线法”避免了一个可能出现的情况，即多个片段应该连接到同一个胞体面片上。这时因为在实际找寻合适的面片时，一些拓扑上应该连接到胞体网格面片上的片段，连接到了某个之前新形成的网格面片上。

## 实验结果和分析

~~我是一个二货。~~

* + 1. 实验环境和方法

本实验具体的环境与开发语言信息如下所示：

处理器：Intel(R) Core(TM) i5-5200U CPU @ 2.20GHz

内存RAM：4.00GB

系统类型：64位操作系统，基于x64的处理器

开发语言：Microsoft Visual C++

程序库：GLFW，GLEW，GLM，OpenGL

首先，从SWC文本文件中读取数据，经过预处理得到上述神经突片段，然后，为每个神经突片段计算表面网格。最后，将所有的片段拼接在一起形成一个完整的神经元的表面网格模型。由于OpenGL直接支持绘制三角形面片，因此，最终的网格模型的面片是以三角形面片为基本组成单位。我们使用一个顶点列表和一个面片列表存储表达网格模型的数据。列表使用C++标准模板库的vector容器实现。顶点列表存放所有顶点的坐标，每个元素表示一个顶点在空间中的位置；面片列表存放构成网格面片的顶点在顶点列表中的索引序号，每三个索引序号表示的顶点构成一个三角网格面片。

将表达网格模型的顶点列表和面片列表传递给OpenGL，OpenGL遵循其渲染流程，按照面片列表中的顶点索引，绘制每一个三角形面片，最终输出整个表面网格模型。

* + 1. 实验结果和分析

图3.X给出了小鼠脑干细胞NMO\_45927的实验结果。从图中可以直观的看到算法所生成的表面网格符合原数据的骨架表示显示出的拓扑结构。

## 小结

本章，我们首先介绍了神经元的基本概念和实验使用的数据来源和含义。然后，介绍了表面网格生成算法的原理。最后，介绍了实验结果以及分析。

第四章 手部姿态动态跟踪系统

从几何造型角度来看，在第三章中重建出的神经元表面薄膜的网格表示是比较粗糙的，这在一些需要XX的场景中是

## 粗糙网格局部优化

考虑到与G代码的兼容性，以及当今绝大部分数控机床… …

## 网格细分

… …。CVS能有效的将G代码转换为CNCS代码。… …

## 数据结构

在进行网格细分时。

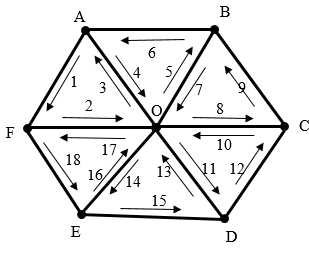
4.3.1 几何形体表示

在第三章中得到了一个表示单个神经元表面的网格模型，并使用一个顶点列表存储构成网格的所有顶点坐标，使用一个面列表表示构成每个网格面片的顶点在顶点列表中的索引，然后将数据传递给OpenGL的数据缓冲，在场景中绘制出来。这种组织数据的方式能够简单直接地在计算机中表示出网格，但是不足以进行更复杂的操作，如网格查询和处理。然而，曲面细分不仅要求增加新顶点，还要求能够将新、旧顶点合并后重新进行拓扑排列，因此，我们需要采用更好的数据结构来描述细分前的模型。

所采用的数据结构不能够只有顶点的位置，而是要表示出网格的拓扑结构，即曲面是如何表达的。这就需要建立起构成网格面片的顶点、边以及该面片之间的关联。一般来讲，这种关联需要方便、快速地获得以下的拓扑信息：哪些面（或边）使用了这个顶点；哪些面使用了这条边；这个面由哪些边构成；哪些面和这个面相邻接等等。满足了这些条件，就可以建立起顶点与边、边与边和边与面之间的关系。半边结构就是一种用于网格数据结构的表达方式。

4.3.2 半边数据结构

半边数据结构最大的特点是以半边为基本单位， 所谓“半边”，就是将一条边看作是两个方向相反的半边构成。除了边界之外，当前边被两个面共有，则该边的两条半边各自属于一个面，而边界边的一条半边为闲置状态。每一条半边只存储起点，这样两条半边就能够表示一条边的两个端点。在对网格进行查询时，只需要沿着半边的顺序，就能够很方便的搜索整个网格。如图X所示，从顶点A开始，可以沿着半边路径1-18-15找到顶点D；从任意半边开始，可以遍历整个网格，如1-2-17-18-16-14-15-13-11-12-10-8-9-7-5-6-4-3。有关半边数据结构的更详细介绍可以在[x]中找到。



需要注意的是，半边数据结构表达的网格需要是流形结构。在计算机图形学中，流形是一种几何模型表面，即二维流形。如果网格的每条边至多被两个面片共用，则该网格就是流形网格，否则就是非流形网格。

第五章 结论和展望

## 流程的制定

预打算用一五轴联动的加工中心加工一螺旋桨… …

## CNCS的应用

改零件的CNCS源代码为：

… …

## CVS的应用

原NC代码为：

… …

## 结论

经过两次加工所得零件… …

第六章 结论与展望

## 6.1结论

本文以一五轴联动的加工中心为研究对象，… …等内容进行了深入的研究，并在有关虚拟加工理论方面进行了一些探索性的工作。主要取得了如下成果：

1. … …；

… …

## 6.2展望

通过本文… …的归纳总结，作者认为在该研究领域还存在以下几方面值得进一步扩充和深入：

1. 该系统与现实CNC系统的真实结合应用；
2. … …；

… …

参考文献

1. Schmitz, B., Virtual Reality: On the Brink of Greatness [J]. Computer Aided Engineering, Vol. 12, No. 4, 1993, pp.26~32
2. Jayaram, S., Connacher, H.I., and Lyons, K.W., Virtual Assembly Using Virtual Reality Techniques [J]. Computer Aided Design, Vol. 29, No. 8, 1997, pp. 575~584
3. Jung, B., Hoffhenke, M., and Wachsmuth, I., Virtual Assembly With Construction Kits [M]. Proceedings of 1997 ASME Design Engineering Technical Conference, September 14-17, 1997, Sacramento, DETC97/DFM-4363
4. 张茂军. 虚拟现实系统[M]. 北京：科学出版社，2001.9：114-169

… …

作者在攻读X士学位期间公开发表的论文

【1】. 黄延军， 戴春祥，胡庆夕，方明伦. 全国高等学校制造自动化研究会第11届学术年会论文集[C]. 上海：上海大学出版社，2004.8：81-83

【2】. … …

… …

作者在攻读X士学位期间所作的项目

【1】. ××省“十五”制造业信息化工程重大科技攻关项目“基于支持智能化×××设计的创新开发平台”

【2】. …　…

…　…

致 谢

本文是在导师张景峤教授的悉心指导下完成的。承蒙张景峤老师的亲切关怀和精心指导，虽然有繁忙的工作，但仍抽出时间给予我学术上的指导和帮助，特别是给我提供了良好的学习环境，使我从中获益不浅。张景峤老师对学生认真负责的态度、严谨的科学研究方法、敏锐的学术洞察力、勤勉的工作作风以及勇于创新、勇于开拓的精神是我永远学习的榜样。在此，谨向张景峤老师致以深深的敬意和由衷的感谢。

还要感谢我的父母，他们在生活上给予我很大的支持和鼓励，是他们给予我努力学习的信心和力量。

最后，感谢所有关心我、支持我和帮助过我的同学、朋友、老师和亲人。在这里，我仅用一句话来表明我无法言语的心情：感谢你们！