

学校代码	10699
分 类 号	TP391
密 级	
学 号	

题目 基于姿态的全手行为跟踪  
与手势交互技术

作者 XXX

学 科 专 业 计算机科学与技术  
指 导 教 师 XXX 教授  
培 养 单 位 计算机学院  
申 请 日 期 2024 年 1 月



西北工业大学

硕士学位论文

题目：基于姿态的全手行为跟踪  
与手势交互技术

学科专业：计算机科学与技术

作者：XXX

指导教师：XXX

2024 年 1 月

**Title: Attitude-based Whole-Hand Behavior Tracking  
and Gesture Interaction Technique**

**By  
XXX  
Under the Supervision of Professor  
XXX**

A Dissertation Submitted to  
Northwestern Polytechnical University

In Partial Fulfillment of The Requirement  
For The Degree of  
Master of Computer science and technology

Xi'an P. R. China  
January 2024





## 摘 要

全手行为跟踪是实现手势交互的重要技术之一，在人机交互的诸多场景中展现了较高的应用价值。目前手势交互领域仍面临交互指令误识别、可穿戴设备错位导致的误差等问题。因此，本文通过开发手部约束特性优化全手行为跟踪方法，研究震颤信息区分交互意图，对空中手势交互领域具有理论与应用价值。

本文主要工作如下：

1) 研究了一种全手行为跟踪优化方法。首先，本文依据传感器姿态估计方法，构建全手行为跟踪算法。其次，开发手部约束特性，将手部约束作为先验知识融合进全手行为跟踪算法中，作为全手行为跟踪优化算法。最后，通过实验验证在对指间关节姿态求解中优化算法相较于未优化算法的  $q_3(F_{1,33} = 396.4, p < 0.001, \eta^2 = 0.46)$  和  $q_4(F_{1,33} = 396.3, p < 0.001, \eta^2 = 0.46)$  误差存在显著差异，表明姿态优化方法结果更加贴近真实手部姿态。

2) 研究了一种基于姿态的手势集合与手势交互实现方法，分析基于震颤的交互意图识别方法。首先，基于以用户为中心的设计方法研究了一种基于姿态的手势集合。其次，分析了坐标映射与手势实现方法。最后，根据用户行为中的震颤特征以识别交互意图，实验表明通过震颤特征区分用户行为有无交互意图具有 97% 的准确率。

3) 设计并实现了基于姿态的空中手势交互原型系统 iGestureGlove。在该交互系统中，首先针对该空中手势交互系统设计模块化数据手套，为系统提供硬件支持。其次基于空中手势交互的基本功能构建交互任务。最后展示了原型系统的实现与应用。

本研究面向基于姿态的全手行为跟踪与手势交互技术，通过开发手部约束特性以提高姿态跟踪的准确度，所用方法可扩展到更多的约束维度，最终提高姿态估计方法的性能和适用性。本研究分析了行为震颤与交互意图的关系，通过用户行为获取更多维度的信息，可以提高空中手势交互的自然性，这对于全手行为跟踪技术以及空中手势交互技术具有参考意义。

**关键词：** 人机交互；手势交互；手部姿态；行为跟踪；交互意图

## Abstract

Whole-hand behavior tracking is one of the important technologies to realize hand gesture interaction, and it shows high application value in many scenes of human-computer interaction. At present, there are still some problems in the field of gesture interaction, such as misrecognition of interaction instructions and error caused by the misplacement of wearable devices. Therefore, this paper studies the interaction intention of tremor information by developing the method of hand constraint to optimize the whole hand behavior tracking, which has theoretical and practical value in the field of air gesture interaction.

The main work of this paper is as follows:

1) A whole-hand behavior tracking optimization method is studied. Firstly, the whole-hand behavior tracking algorithm is constructed based on the IMU attitude estimation method. Secondly, the hand constraint is fused into the whole-hand behavior tracking algorithm as a prior knowledge and used as the optimal algorithm of whole-hand behavior tracking. Finally, experimental verification demonstrated a significant difference in errors between the optimized algorithm and the non-optimized algorithm in solving the posture of the interphalangeal joint  $q_3 (F_{1,33} = 396.4, p < 0.001, \eta^2 = 0.46)$  and  $q_4 (F_{1,33} = 396.3, p < 0.001, \eta^2 = 0.46)$ , indicating that the posture optimization method yields results that are closer to the actual hand posture. This further confirms the superior accuracy of the posture optimization method in approximating the real hand posture.

2) A gesture-based interactive implementation method of gesture set and gesture is studied, and a tremor-based interactive intention recognition method is proposed. Firstly, a gesture set based on the user-centered design method is studied, and then, the method of interactive coordinate mapping and gesture implementation are analyzed, the experimental results show that 97% accuracy can be achieved by extracting user tremor information to identify user interaction intention behavior.

3) A prototype system for aerial gesture interaction, called iGestureGlove, has been designed and implemented. In this interaction system, a modular data glove was designed specifically for the aerial gesture interaction system to provide hardware support. Basic functions for aerial gesture interaction were then constructed to facilitate interaction tasks. Finally, the implementation and application of the prototype system were demonstrated.

This research aims at the gesture-based whole-hand behavior tracking and gesture interaction technology. By developing the hand constraint feature to improve the accuracy of



gesture tracking, the method used in this research can be extended to more constraint dimensions, finally, the performance and applicability of the attitude estimation method are improved, and the relationship between behavior tremor and interaction intention is analyzed, it can improve the naturalness of air gesture interaction, which has reference significance for the whole hand behavior tracking technology and air gesture interaction technology.

**Key words:** Human-Computer interaction; Gesture recognition; Hand attitude; Motion tracking; Interaction intention

# 目 录

摘 要 .....	I
Abstract .....	II
第 1 章 绪论 .....	1
1.1 研究背景与意义 .....	1
1.2 国内外研究现状 .....	2
1.3 主要研究内容 .....	5
1.4 论文结构 .....	6
第 2 章 手部跟踪与交互方法相关研究 .....	7
2.1 手部跟踪相关研究 .....	7
2.1.1 基于磁感应的手部跟踪 .....	7
2.1.2 基于计算机视觉的手部跟踪 .....	8
2.1.3 基于惯性传感器的手部跟踪 .....	9
2.2 手势识别相关研究 .....	10
2.2.1 基于原始数据的手势识别 .....	10
2.2.2 基于手部姿态的手势识别 .....	11
2.3 本章小结 .....	12
第 3 章 基于姿态的全手行为跟踪 .....	13
3.1 姿态表示与关节特性 .....	13
3.1.1 姿态描述方法 .....	13
3.1.2 手部关节特性 .....	16
3.2 全手行为跟踪方法 .....	18
3.2.1 IMU 姿态求解方法 .....	18
3.2.2 全手行为跟踪方法 .....	20
3.3 全手跟踪优化方法 .....	21
3.3.1 全手行为跟踪的误差分析 .....	22
3.3.2 构建手部非线性约束模型 .....	23
3.3.3 非线性约束下的姿态求解 .....	26
3.3.4 实验分析 .....	27
3.4 本章小结 .....	30
第 4 章 基于姿态的手势交互技术 .....	31
4.1 基于姿态的平面交互手势设计 .....	31
4.1.1 平面交互基本功能 .....	32
4.1.2 用户设计手势实验 .....	33

4.2 平面空间下手势交互实现方法 .....	37
4.2.1 手部姿态与平面空间坐标映射 .....	38
4.2.2 基于手部姿态的交互指令实现 .....	40
4.3 基于姿态的交互意图识别方法 .....	41
4.3.1 手部行为震颤与交互意图 .....	42
4.3.2 交互中手部震颤提取方法 .....	43
4.3.3 基于震颤的交互意图识别 .....	46
4.3.4 实验分析 .....	50
4.4 本章小结 .....	58
第 5 章 iGestureGlove 原型系统设计与实现 .....	59
5.1 iGestureGlove 的总体设计方案 .....	59
5.1.1 原型系统硬件设计 .....	59
5.1.2 原型系统软件结构 .....	61
5.2 iGestureGlove 原型系统实现与应用 .....	62
5.3 本章小结 .....	64
第 6 章 总结与展望 .....	65
6.1 总结 .....	65
6.2 展望 .....	65
参考文献 .....	67
致 谢 .....	73
攻读硕士学位期间发表的学术论文和参加科研情况 .....	74



# 第1章 绪论

## 1.1 研究背景与意义

随着信息技术的快速发展，人机交互方式日益丰富化和智能化。手势交互作为一种自然、直观的交互方式，越来越多地应用到日常生活中，与语音<sup>[1]</sup>、眼动<sup>[2]</sup>、键鼠<sup>[3]</sup>、手柄<sup>[4]</sup>等交互方式相比，手势交互具有操作空间大、表达能力强的特点，是实现人机自然交互的重要方式之一。作为人机交互领域中自然、灵活、高效的交互模式之一，手部交互能够提供人手行为的直接映射，降低交互过程中的操作复杂度，提高交互效率，从而进一步地提高用户的交互体验。

手势交互已经广泛应用于日常生活的各个领域，而随着这些应用交互行为复杂度的提高，对手部运动跟踪的精确度要求也越来越高。例如，基于手势交互的远程遥控<sup>[5]</sup>系统往往需要实现复杂的精细化控制，如远程手术中通过机械臂进行伤口缝合。医疗康复评估和治疗中<sup>[6]</sup>，高精度的手部运动捕捉可以帮助医生更准确地判断患者的康复状况并制定后续方案，也可以使患者通过运动反馈来纠正错误动作，避免二次伤害。手部跟踪技术的不断进步使这些医疗评估和治疗活动量化和规范化，从而更加精准高效。总体来说，这些应用对手势交互精确度提出了更高要求，如何继续提高手部行为跟踪的准确性仍是该领域的关键研究方向。

在手势交互的应用领域，往往在不同的交互环境中，会有着独立的手势交互指令与手势集，这意味着用户在使用不同的交互设备时，不合理的手势集合与指令集和会加重用户的负担。例如在手机的智慧交互领域，通过单手的旋转和抓握实现了翻页、截屏等操作。在智慧驾驶领域<sup>[7]</sup>，通过手势实现接听电话、切换歌曲、控制音量、快捷导航等操作，在熟悉操作手势之后，会使驾驶过程中的操作更加自然便捷，也使得驾驶员的驾驶体验更加舒适愉悦，但是复杂且难记的操作指令往往会适得其反。在智慧课堂领域中，通过适配且符合场景的手势控制呈现互动式、情景式教学课堂，不仅可以极大地激发学生的积极性，更能够加深学生对知识的理解和记忆，但重复且冗余的手势则适得其反，增加教师的疲劳、干扰学生学习知识。此外，针对不同交互任务和场景，手势功能范围存在差异。因此如何根据交互任务和环境特征，定义和选择合适的手势集合是手势交互系统中需要考虑的环节，同时提高交互过程中的流畅度，通过交互在真实环境与虚拟环境中流畅的完成交互目的，也是手势交互应用领域中所面临的挑战。

综上所述，日益增长的应用场景和需求，以及提升用户体验的要求，都使手部跟踪技术在跟踪效果和交互体验面临新的挑战。因此研究更高精度的全手行为跟踪技术和自然的手势交互技术，具有重要的研究和应用价值。

## 1.2 国内外研究现状

目前,手部交互可以分为手部跟踪和手势识别两部分,手部跟踪主要目的在于获取手部的空间状态,而手势识别的主要目的则在于将手部信息转化为指令信息。目前手部跟踪的方法多种多样,有基于视觉、Wi-Fi、惯性传感器和多传感器融合的手部跟踪。手部跟踪可以获取手部运动中的更多细节,能够提供更加灵活和自主的交互,在一些注重手部交互细节的场景中具有更多优势。由于手部拥有多自由度,其空间姿态信息相对复杂,因此手部跟踪具有非常重要的研究价值。

基于视觉的手部跟踪,通常使用相机捕捉手部图像,然后利用机器学习和优化算法识别手部特征,达到手部跟踪的目的,其中单目、双目和深度摄像机是基于视觉的姿态检测系统中使用的三种主要摄像机类型。基于视觉的方法进行手部跟踪具有无接触、非侵入性,使用方便。计算效率高,许多算法已经实现了实时追踪和识别等优点,现阶段已经有许多商业化设备提供姿态估计,微软的第一代 Kinect 是一个深度相机,其结合了 SDK 库来监测人类关节的骨骼,作为手势识别研究的基础。Hisham 等人<sup>[8]</sup>为 Kinect 引入了一个动态的阿拉伯手语识别系统,将决策树与贝叶斯分类器相结合用于手势识别,并采用 AdaBoost 方法对系统进行改进,识别率达到 93.7%。Xi 等人<sup>[9]</sup>提出一种基于 Kinect 的实时手部跟踪方法,在手部区域提取过程中,使用一种级联结构和递归连接组件算法,提高计算效率并保留三维空间中的连接关系,在采样率大于 45 帧的条件下平均误差在 23 毫米。Mauro 等人<sup>[10]</sup>对新版 Kinect 进行评估,身体主要关节的平均均方根误差约为 0.47,角度平均误差在 5-15 度之间,同时指出在缓慢运动时估计姿态有更好的表现。Elyoenai 等人<sup>[11]</sup>使用 Leap Motion 控制器获取手部姿态信息进行空中手势签名验证,最终识别结果的等错误率分别为 1.20%。



图 1-1 Leap Motion 与 Kinect 设备

基于视觉的手部跟踪技术除了使用特殊摄像头外,使用普通摄像头进行手部跟踪也得到了越来越广泛的关注。Cai 等人<sup>[12]</sup>介绍了一种使用合成数据和弱标记的 RGB 图像进行 3D 手部姿态估计的方法,从单目 RGB 图像推断空间手部姿态,将带有深度正则化器辅助标记 RGB 数据集用作手部姿态,预测并推断空间手部关节位置。Danilo 等人<sup>[13]</sup>使用多任务语义特征提取器从 RGB 图像中生成平面热图和手的轮廓,利用弱透视投影和神经网络估计手部关节姿态。Li 等人<sup>[14]</sup>提出了一种基于静态 RGB 图像进行手势联合

定位和识别的高效深度注意技术, 该技术根据软注意力策略训练了卷积神经网络(CNN, Convolutional Neural Network)模型, 实现了手部的自动定位和手势识别。Pinto 等人<sup>[15]</sup>提出了一种基于 CNN 的高效手势识别方法, 这些步骤包括分割、多边形逼近、轮廓生成和形态学滤波器的应用, 有利于增强特征提取, 最后使用不同的 CNN 进行测试比较, 在预设数据集上识别准确率为 96%。Chen 等人<sup>[16]</sup>提出了一种新的深度学习模型, 用于解决基于 RGB 图像的 3D 手部姿态估计, 其引入了一个层次化的 2D 到 3D lifting 模块, 通过首先估计手部关节的 2D 坐标, 然后将其提升到 3D, 在 RHD 手势数据集的误差为 10 毫米。

基于视觉的手部跟踪研究, 除了依靠外部环境的视觉设备进行手部跟踪外, 也有学者尝试将放置在场地的视觉设备转移到用户身上, 实现基于可穿戴视觉设备的手部跟踪。Pail 等人<sup>[17]</sup>提出了一个称为 HOOV 的手部动作跟踪系统, 系统使用佩戴在手腕的惯性传感器, 可以持续估计用户手部在 3D 空间中的位置, 实现屏幕外交互, 平均跟踪误差为 7.7 厘米。使用视觉设备进行手部跟踪的方式尽管已经取得了很高的准确度, 但仍存在多种局限性, 比如容易受到环境光、遮挡、图像失真等影响, 并且隐私保护也是不可避免的问题。

基于无线信号的手部跟踪, 通过无线信号的发射和接收时间差测量人手各部分的距离信息, 从而实现手部跟踪。近年来 Wi-Fi 设备现在已经出现在许多室内环境中, 2000 年, Bahl 等人<sup>[18]</sup>首次将 Wi-Fi 信号用于人体感知, 他们提出了一种基于 Wi-Fi 的室内定位系统。Pu 等人<sup>[19]</sup>提出了一种无线捕捉系统, 使用多普勒频移作为识别 9 种手势的特征, 对于识别预设手势的准确率为 94%。Zhang 等人<sup>[20]</sup>提出了与手势特征相关的动态相位矢量和旋转变量特征, 并采用分层感知框架构建了实时的手写手势跟踪系统, 在动态与静态环境下的准确率超过 95%。Han 等人<sup>[21]</sup>使用三个普通无线设备, 采用多信号分类算法来估计动态路径的手部角度, 并通过三角测量方法来定位手的初始位置, 最终平均跟踪误差为 1.3cm, 手写识别的准确率为 94%。Mudra 等人<sup>[22]</sup>通过利用不同位置的天线之间的信号差异开发了一种基于 Wi-Fi 的毫米级手势检测方法, 手势识别的准确率为 96%。Fall 等人<sup>[23]</sup>应用随机森林方法和 SVM 分类器对各种人类活动进行分类并实现跌倒检测, 平均错误率为 18%。Fahd 等人<sup>[24]</sup>提出了一种基于 Wi-Fi 信号的无线传感方法, 利用信号处理技术实现对手部活动的高准确度识别, 利用深度学习算法使分类的准确率达到了 97.5%。结合上述研究内容, 可以发现, 基于无线信号的手部跟踪不需要手接触任何传感器, 具有良好的隐私性等优点, 但也具有容易受到多手指的互相干扰、追踪效果受距离影响较大等特点。

基于惯性传感器 (IMU, Inertial Measurement Unit) 的手部跟踪, 利用了惯性传感器集成的加速度计、陀螺仪和磁力计的测量信息, 融合计算传感器的姿态, 进而实现手部跟踪。Kortier 等人<sup>[25]</sup>提出使用安装多个惯性传感器的数据手套获取手部的运动信息, 在

运动过程中测量所有手指和关节的姿态，从而实现手部的完整三维重建，但是该设备使用 USB 接口传输数据，无法实现移动环境下的手势交互。清华大学的 Fang 等人<sup>[26]</sup>深入研究了基于数据手套的手部跟踪和交互技术，使用传感器融合算法估计关节的姿态，并利用正向运动学从关节姿态推算指尖位置，从而实现对手部的位姿跟踪。Lin 等人<sup>[27][28]</sup>则针对模块化的数据手套展开研究，利用数据手套上的惯性传感器获取手部各关节的运动信息，再通过融合算法测量用户手部各关节的运动范围，从而跟踪用户的手部关节姿态，其模块化的设计使得数据手套的灵活性和可维护性得到提高。Chang 等人<sup>[29]</sup>提出一种传感器融合算法，在融合数据手套的加速度、磁场强度和角速度数据基础上，利用反馈思想不断修正手部跟踪的方向，并搭建电机平台进行验证，测得平均误差小于 3 度。总的来说，惯性传感器适用于移动环境下的手部追踪，通过加速度进行空间定位时则具有较高的误差，但惯性传感器在实现姿态求解方面具备很高的准确度，因此可将惯性传感器用于手部姿态求解的研究中。

基于多传感器融合的手部跟踪，往往通过多个数据源拟合出最优姿态解，以弥补各自的缺点。现有通过基于佩戴惯性传感器的数据手套估计手部姿态的方法仍然存在着误差漂移问题和由于碰撞等原因导致的姿态明显不符合手部结构问题，为了提高手部姿态估计的性能，有研究者通过多传感器融合的方式提高姿态估计的精度。Jiang 等人<sup>[30]</sup>将数据手套与深度相机配合使用，将基于可穿戴的手部姿态估计方法与基于模型的手部姿态估计处理方法相结合，提高了手部姿态估计的稳健性和准确性。Park 等人<sup>[31]</sup>将佩戴于手部的陀螺仪与佩戴于头部的深度摄像头结合，利用陀螺仪良好的动态性能，弥补了手在高速运动时深度图像失真导致手部姿态估计准确性大大降低的问题。Liu 等人<sup>[32]</sup>通过建立描述手指间运动动力学的手掌模型，提高了手指跟踪精度，结合深度相机 Kinect 对手掌位置进行跟踪，进一步减小了使用传感器进行位置估计的误差。基于多种传感器的手部跟踪。Li 等人<sup>[33]</sup>通过融合传感器的姿态信息和磁感应强度信息，解算得手部的位姿信息，探索了磁感应手势跟踪的方法。也有 Liu<sup>[34]</sup>等人探究脑电信号与惯性传感器融合的方式进行手势跟踪，利用脑电信号与惯性传感器互补，得到更优的姿态解。将惯性传感器与其它传感器结合，虽然提高了手部姿态估计的稳定性和准确性，但也增加了用户的负担，提高了功耗和成本。

有研究者考虑到人体手部约束属于身体属性之一，将生理约束融入到手部姿态估计中，通过针对手部不同的生理学知识进行处理，设计相关手部模型<sup>[35][36]</sup>减少姿态误差。Wen 等人<sup>[37]</sup>提出了融合偏差减少传感器误差的思想，通过对传感器测量误差的建模处理，提高了手部姿态解算的性能，其次基于手部约束建立姿态求解模型，但多个求解模型形成嵌套关系，增加了系统求解时间。Liu 等人<sup>[32]</sup>通过建立描述手指间运动动力学的手掌模型，结合视觉识别和定位方法，补偿由传感器精确导致的手掌位置估计误差，提高了手指跟踪精度，但并未深入探究关节约束特点。Li 等人<sup>[38]</sup>提出了将生理学手部知识



融合进三维手模型之中,将图像中的肌肉、骨骼和皮肤与模型通过迭代形状配准和参数学习进行对应匹配,以保证得到的三维手模型的运动范围与姿态符合物理手部结构。手部生理学约束<sup>[39][40]</sup>包含关节的运动角度存在的约束关系,手部各关节之间不是相互独立的,并且各关节拥有的自由度不同,某个关节的动作可能会引起其他关节做出相应反应,关节之间存在着约束和依赖关系,同样单个关节的活动范围也是受限的,这些关系导致手部各组成部分密不可分,共同形成了手这一整体,本文发掘手部约束,对关节约束转化为状态约束,得到符合物理手部结构的手部姿态。

在手势交互领域中,为了进一步使交互系统提供更流畅的交互体验,在交互过程中获取并识别用户的交互意图成为了一种研究方向。交互意图是由人类活动意识导致大脑的神经细胞产生有节奏的电信号<sup>[41]</sup>产生的,可以利用放置在头皮区域的电极检测反映人类意识活动的脑电信号,进而实现控制假肢等操作。用户行为中,动作意图识别主要分为两个步骤,特征提取和分类识别,提取的特征主要包括时域<sup>[42]</sup>、频域<sup>[43]</sup>等特征<sup>[44]</sup>,在识别不同动作的过程中,单个特征的分类效果可能会有所不同。影响动作识别的另一个因素是机器学习算法的选择。目前常用的识别动作意图的算法包括支持向量机(SVM)、决策树和神经网络等。Mishchenko 等人<sup>[45]</sup>开发了一种基于脑电图的非侵入性脑机接口系统,该系统通过脑电数据检测用户的心理意图,平均准确率达到 80%。Zhong 等<sup>[46]</sup>利用时频信息融合方法获得表面肌电信号的时频特征,并引入深度信念网络来识别上肢的执行动作。Ning 等人<sup>[47]</sup>提取时域和频域的表面肌电信号,并采用优化的支持向量机来识别上肢运动。Hou 等人<sup>[48]</sup>利用惯性传感器记录被测物体的运动轨迹,并利用长短期记忆来识别身体活动。Sun 等人<sup>[49]</sup>利用惯性传感器获取假肢膝关节的姿态数据,并利用 BP 神经网络实现假肢膝关节的自适应速度控制。上述研究者通过识别交互意图预测或补全用户行为,进而提高模型识别的准确率。也有研究者<sup>[50]</sup>利用震颤作为交互输入,利用线性加速度计检测人体生理震颤作为移动设备新型输入的方法,这使得设备能够感知是否被握在人手中,以及握持的姿势。目前,利用可穿戴设备的优势,从用户行为中提取震颤信息用于交互意图的识别,正逐渐成为探究自然交互的一种方法。

根据现有研究,本文结合目前的手部跟踪应用及需求,对手部跟踪和交互技术展开研究,旨在实现更精确的动作捕捉,并根据本文的交互环境探索自然的手部交互方式。

### 1.3 主要研究内容

本文的研究内容主要分为三个部分:

1) 本研究通过开发手部约束特性,提出一种全手行为跟踪优化算法,降低了全手行为跟踪的误差。探究手部姿态的常见构型和表示方法,建立以四元数为基础的算法模型。开发手部约束特性,通过将融合约束对全手行为跟踪算法进行优化,降低可穿戴设备与手部错位的误差,得到符合生理结构的手部姿态,提高手部姿态估计的准确度。

2) 本研究建立了一种用户参与设计的空中交互手势集合,降低用户的学习成本,并

研究了一种基于决策树的手势交互实现方法。最后研究用户行为中震颤特征与交互意图的联系，通过采集用户运动信息，提取震颤特征以区分用户行为是否具有交互意图，设计用户实验并验证了该方法具有可行性。该方法能使交互系统理解用户交互意图，解释用户交互行为，实现更流畅的手势交互。

3) 设计并实现了手势交互原型系统 iGestureGlove。在该交互原型系统中设计了一种模块化无线传输数据手套，为基于姿态的空中单手手势交互提供硬件支持。结合全手行为跟踪优化算法，给出原型系统的实现与应用场景，实现了自然的手势交互。

## 1.4 论文结构

论文的章节结构安排如下：

第一章是绪论部分，从本文的选题背景和研究意义展开叙述，随后针对手部跟踪的国内外研究现状做出简要描述，最后提出本文的主要研究内容，并对论文的总体结构进行介绍。

第二章介绍手部跟踪及交互技术的相关研究，从全手行为跟踪技术和手势交互技术两方面展开。全手行为跟踪技术中介绍了基于电磁波、计算机视觉、惯性传感器的手部跟踪技术。手势交互技术中分析了基于原始数据的手势识别和基于姿态的手势识别。

第三章主要研究融合手部约束的全手行为跟踪优化方法。首先建立了基于四元数的手部模型，为了降低可穿戴设备与手部运动的相对位移误差，将基于人手的关节约束建模为基于四元数姿态的数值约束，并将其作为先验知识融合进姿态求解算法中，得到符合真实人手物理约束条件的手部姿态，最后通过设计用户实验，验证该优化方法相较于未融合约束方法提高了手部跟踪的准确率。

第四章研究了一种基于姿态的手势集合与交互技术，并探讨了一种交互意图识别方法。首先采用以用户为中心的设计方法设计基于姿态的手势集，向用户解释手部姿态含义并采用大声思考法记录用户设计想法，建立一种以用户为中心的手势集合。根据该手势集合探索了一种坐标映射方法和手势识别方法，以实现长短距离映射以及交互指令映射。最后提出了一种用户交互意图识别方法，通过提取震颤特征识别用户行为，设计实验并验证了该方法的可行性。

第五章面向空中手势交互领域，设计并实现了一个空中手势交互原型系统 iGestureGlove，介绍了系统硬件和软件两部分。最后展示了 iGestureGlove 原型系统的实现与应用场景。

第六章主要对全文的研究内容和系统实现做出总结，并展望了今后的研究方向。

## 第2章 手部跟踪与交互方法相关研究

本章介绍了当前手部跟踪算法的常见实现方法，包括基于磁感应、计算机视觉、惯性传感器的手部跟踪，分析了三种方法各自的优缺点。介绍并分析了基于原始数据和手部姿态的手势识别方法。

### 2.1 手部跟踪相关研究

#### 2.1.1 基于磁感应的手部跟踪

基于磁感应的手势跟踪是利用佩戴在手指及手腕处的磁传感器捕捉手部运动，获取手势形状、方向以及动作变化信息。

在基于磁感应追踪的手势识别中，往往通过穿戴磁感应设备进行手部跟踪，由三轴磁感应传感器、信号放大及滤波电路、微控制器等组成。用户首先需要在手部关键位置佩戴微型磁传感模块，通过微处理器解决方位标定和时间同步问题，确保模块之间的协调运作。通过磁场发生装置生成稳定而精准的参考磁场，手指运动干扰该磁场，磁感应传感器探测微小磁通密度的变化，并引起输出电压值的变化，从而测算各模块在参考坐标系下的实时空间位置和方向。最后，读取不同时间段的姿态数据，可以构建动态手势的骨骼运动学模型。

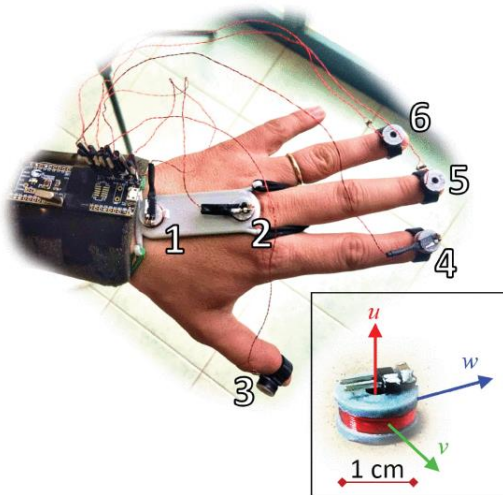


图 2-1 磁感应跟踪系统佩戴布局<sup>[53]</sup>

目前使用磁感应进行手部跟踪有了广泛的研究，Phil 等人<sup>[51]</sup>提出了通过多传感器共同检测同一目标点的方法，预设传感器已知的姿态偏差和位置偏差，为消除传感器姿态影响提供了一种有效的方法，然而未解决磁场空间姿态的影响，在实验中将其假设为固定值。Pasku 等人<sup>[52]</sup>采用三轴传感器线圈，利用谐振特性降低了电流和功耗需求，并具有较高的定位精度。Santoni 等人<sup>[53]</sup>提出的 MagIK 系统是一种基于手部运动学模型的磁传感器手势跟踪系统，该模型将手部分割成多个刚体段，每个刚体段均安装磁传感器以

捕捉运动，其次基于优化算法校准传感器参数并提供高精度的手部定位，在手势跟踪中测得误差小于 10 毫米。

基于磁感应的手部跟踪技术具有许多优点，例如在短距离内手部跟踪具有较高精度和较低延迟。磁感应技术不受视线遮挡的影响，适用于复杂环境下的跟踪需求。

然而该技术也存在一些缺点，例如当受到金属物体干扰或磁场变化时导致姿态估计出现误差。由于需要在用户手部放置特殊的磁性传感器，长时间的手势交互可能会导致用户疲劳。磁场强度受距离影响较大，随着与磁场源距离增大导致姿态误差也增大。

### 2.1.2 基于计算机视觉的手部跟踪

基于计算机视觉的手势跟踪技术利用图像处理和机器学习算法捕获和分析手部运动，实现手势跟踪。

首先系统需要采集手部图像或视频作为算法分析的输入源，高质量的图像采集对整个系统识别性能至关重要，需要捕捉清晰完整的手部视角。其次，基于视频或图像进行手部检测与分割，运用肤色检测、背景差分等图像处理技术区分前景手部和背景元素，以定位手部区域。正确的手部定位是手势特征提取的先决条件。然后在分割后的手部图像上进行轮廓描边、关键点提取以获取手势形状和姿态的数据。例如，检测手指尖点来判断指向，通过掌心 and 指骨的轮廓信息识别手掌的方向和形态，这些都是重要的手势特征。基于上一步得到的特征数据，建立对应的机器学习模型以实现手部形状和动作的自动分类与识别。通常需要事先从样本图像中训练模型，让其学习不同手势的特征判别方法。最后，系统会连续跟踪和解析视频序列，以输出手势的时间序列信息和解释手势所表示的含义并转换为计算机指令或相应交互。

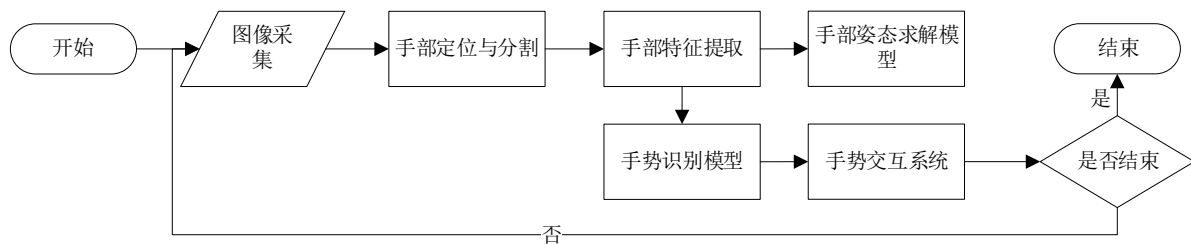


图 2-2 基于计算机视觉的交互流程图

基于计算机视觉的手部跟踪技术具有许多优点，例如通过摄像头捕捉手部运动，实现无需额外设备的自然交互，提高了用户体验。由于普通摄像头的普及，计算机视觉系统相对较低的硬件成本。随着机器学习与深度学习的发展，基于计算机视觉的手部跟踪技术日益成熟，手势识别技术具有较高的精度。

然而该技术也存在一些不足，首先复杂的图像处理要求计算机具有较高的性能，也受限于摄像头的采样率，导致交互过程中存在不流畅等问题。其次基于视觉的手势跟踪技术容易受到环境光的影响，如在不同的光照条件下同一个识别模型会出现不同的识别

精度。基于计算机视觉的手势跟踪技术还往往受限于摄像头的拍摄范围，当距离摄像头较远时出现模型错误识别的现象。用户的个体差异也会影响识别效果，例如用户的手掌大小，肤色等。综合而言，尽管计算机视觉技术为手势交互带来了革命性的进步，但仍需要根据交互需求与环境选择。

### 2.1.3 基于惯性传感器的手部跟踪

基于惯性传感器的手势跟踪是通过可穿戴设备采集手部运动数据以实现动作捕捉与手势识别。手势跟踪的第一步是构建手部构型并佩戴 IMU。九轴 IMU 一般包括陀螺仪、加速度计和磁力计，分别测量手部运动的角速度、加速度和磁感应信息。为捕捉整个手部骨骼的全详细运动，需要在手指、手腕等关键点佩戴多个 IMU，然后通过数据融合算法构建手部模型，并利用获取的关节信息通过前向动力学进行手部跟踪。

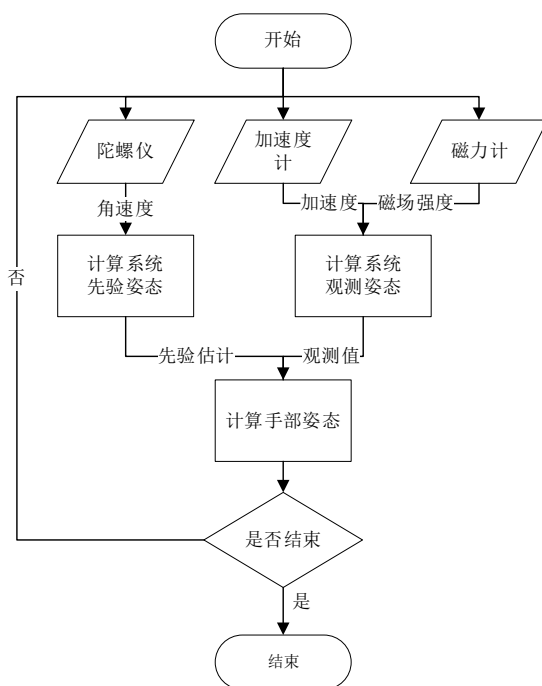


图 2-3 卡尔曼滤波算法流程图

基于惯性传感器的手势跟踪算法一般有互补滤波算法、卡尔曼滤波算法、Mahony 滤波算法和 Madgwick 滤波算法等。Mahony 滤波算法和 Madgwick 算法均是一种基于四元数的滤波方法，用于融合加速度计、陀螺仪和磁力计的数据，该算法相对较轻量，适用于嵌入式系统和实时应用。互补滤波算法是一种简单而有效的滤波方法，通过将加速度计和陀螺仪的数据进行加权平均，以获得更稳定的姿态估计，磁力计的数据可以用于进一步校准姿态，优点为简单、计算效率高，适用于实时应用。卡尔曼滤波系统可以融合加速度计、陀螺仪和磁力计的测量数据，通过对系统状态进行动态估计和修正，以获取最优姿态，卡尔曼滤波系统能够在动态环境中提供良好的姿态估计，对噪声和误差有较强的抑制能力的优点。

基于惯性传感器的手势跟踪技术，其优点包括具有良好的隐私性，不会获取用户声音、人像、环境等信息。数据手套便携易用，这使得手势交互不受空间限制。该技术受环境影响较小，在不同的环境下手势追踪模型往往具有相同的性能。

然而，基于惯性传感器的手势跟踪技术存在一些不足之处。首先，指尖位置的推算依赖于构建的手部模型，因此难以适应不同用户的手部结构差异。其次，多组独立的传感器姿态与实际手相比，缺乏关联性，可以考虑手部肌肉约束建立手部模型，进行全手行为跟踪。此外，用户佩戴可穿戴设备交互过程中，手部与可穿戴设备之间会由于佩戴方式而造成相对位移，增大了手势跟踪的误差。因此，为提升该技术的性能，未来研究可着重解决上述挑战，以实现更为精准、可靠的手势交互体验。

## 2.2 手势识别相关研究

### 2.2.1 基于原始数据的手势识别

基于原始数据的手势识别是指直接利用来自传感器的原始信号信息进行手势分类和识别，而不提取结构化的手势描述符特征。具体来说，采集系统会输出包含手部多维信息的动态序列数据，像从 RGB 或深度摄像头获取的图像帧或者佩戴的电极阵列、超声波雷达接收到的电信号记录。这些数据保留了完整的手势运动细节。然后直接输入神经网络等机器学习模型中，模型会自行学习区分不同手势模式所需的关键特征。通过基于原始数据的方式完成训练，可以获得一个泛化能力好的手势分类或回归系统。

在基于原始数据进行手势识别的方法中，长短时记忆网络（LSTM, Long Short-Term Memory）表现出对手势时序信息进行建模和识别的良好性能。然而，由于 RGB 图像通常具有较高的分辨率，将其直接作为 LSTM 网络的输入并不可行。通常情况下，研究者采用卷积神经网络对 RGB 图像分段进行特征提取，然后将生成的特征序列作为输入传递给 LSTM 网络进行进一步处理。这样的策略在实际应用中旨在有效地捕捉手势的空间和时序特征，提高识别性能。通过这种协同工作的方式，CNN 负责提取静态图像特征，而 LSTM 则专注于处理时序信息，使得整个系统能够更准确地识别和理解手势动作。这种联合使用的方法不仅克服了 RGB 图像输入的挑战，还为手势识别任务提供了更强大的建模能力。

有许多研究者对提出了基于原始数据的手势识别方法，Molchanov 等人<sup>[54]</sup>提出一个基于原始数据的手势识别模型，基于循环三维卷积神经网络的方法，能够从多模态数据中同时进行动态手势的检测和分类，采用连接主义时序分类来训练网络，以从未分割的输入流中预测正在进行手势的类标签。Camgoz 等人<sup>[55]</sup>出了一种名为 Sub Unets 的深度学习方法，还允许在不同相关任务之间隐式进行迁移学习，这种基于原始数据的深度学习模型能够使用更广泛、更多样化的数据源进行识别。Cao 等人<sup>[56]</sup>提出了一种新颖的循环 3D 卷积神经网络，设计了一个具有在邻近时间片之间的循环连接的时空变换器模块，用于解决佩戴可穿戴设备（如 VR/AR 头盔和眼镜）中的手势识别问题。

采用基于原始数据的手势识别方式具有显著的优势，主要体现在避免了人工特征提取和预处理步骤，从而减少了数据加工处理阶段的信息损失，同时保留了手部运动的完整细节，有利于进行差异特征学习。此外，该方法还能够有效实现对多源数据（如视觉、磁信号等）的并行建模，为综合信息的融合提供了便利。

然而，基于原始数据的手势识别方式也存在一些不足。首先，在挖掘原始数据特征的过程中，对样本数量需求较高，通过大量的原始数据才能挖掘出关键的特征信息。其次，该方法需要强大的计算资源支持，提取特征与分类识别的计算需求由模型结构决定。同时，由于其较为简化的特征提取步骤，难以进行对特征的控制变量调整，使得研究人员可能面临较为复杂的黑盒问题，难以深入理解系统内部的运作机制。因此，在采用基于原始数据的手势识别方式时，研究人员需充分利用其优势的同时，认识到其局限性，并寻求相应的应对策略。

### 2.2.2 基于手部姿态的手势识别

基于姿态的手势识别是指通过分析手部骨骼框架的位置和运动情况来判断手势，这类方法模拟了人眼分辨手势的视觉注意机制。首先构建手部的骨骼模型，典型的是使用图像或深度信息预测 2D 或 3D 的关键点坐标，比如手腕、指尖和指节的位置。其次在时间轴上跟踪模型的移动，获取表达手部姿态与运动的状态信息。相较直接应用像素或信号，这种姿态表达的数据量减小了数量级，利于压缩不同状态下手势类型的差异。手部姿态信息往往只包含角度、方向等信息，因此识别模型通常会提取基于姿态的数据特征，例如关键点距离、相对角度以及轨迹形变参数等。最后，将这些特征输入分类器或回归器中，对其进行判别与映射，输出对应的手势类型和参数。

基于手部姿态的手势识别方法是利用手部关键点的信息来进行手势识别。相比于基于 RGB 图像与视频的手势识别方法，手部姿态不受背景信息的影响，能够更好地关注手部的位置与运动信息，是一种识别模型受其他环境影响较低的方法。基于手部姿态的手势识别首先利用手部姿态检测方法获取手部的姿态信息，其次利用传统特征提取方法或深度学习特征提取方法提取手部姿态的特征，最后将提取的特征输入到分类器中进行手势分类。Boulaia 等人<sup>[58]</sup>从手指、掌心和手腕的三维坐标中提取 HIF3D 特征，用以表示手部形状的高阶信息，通过引入时间金字塔来捕获时间信息，最终他们使用支持向量机对提取的特征进行分类。Zhao 等人<sup>[59]</sup>提出了一种基于骨架的动态手势识别方法，该方法提取了四种手部形状特征和一种手部方向特征，然后利用时间金字塔对这些特征进行表示，形成最终的特征向量，最后将该特征向量输入线性支持向量机分类器进行手势识别。

基于手部姿态的手势识别方法往往数据量较少，输入数据只包含手部姿态信息，这使得识别系统具有较低的复杂度。手部姿态相较于原始数据更为直观，方便研究人员总结特征信息，更好的理解识别模型的运作机制。

基于姿态的手势识别技术受姿态求解算法的影响较大，良好的姿态估计与特征选择会降低系统运行的复杂度，但也会造成误差累积与信息损失。因此使用基于姿态的手势识别技术，对手部姿态估计算法提出了较高的要求。

## 2.3 本章小结

本章主要介绍了手部跟踪以及交互技术的相关研究。首先介绍了分别基于磁感应、计算机视觉、惯性传感器的手部跟踪方法，分析其各自优点与缺点。其次介绍了基于原始数据的手势识别技术和基于手部姿态的手势识别技术。依据上述研究，为探究基于惯性传感器的全手行为跟踪技术，通过姿态进行手势识别，为自然手势交互技术提供支持。



## 第3章 基于姿态的全手行为跟踪

本章建立基于四元数的手部姿态求解方法，开发手部生理约束，提出一种全手行为跟踪优化方法，该方法降低了手部与可穿戴设备佩戴错位产生的姿态误差。

在第二章中描述了目前较为常见的手部跟踪的方法，有基于视觉的方法、基于惯性传感器的方法等，本研究选择了由惯性传感器组成的可穿戴设备进行全手姿态跟踪。在使用可穿戴设备进行手势交互时，人手大小的差异、运动习惯的多样性、运动幅度的不一致性均会产生程度不同的佩戴错位。根据分析发现，佩戴错位会出现手部姿态的错误估计，例如通过关节计算手部模型时，出现两个手指模型部分重合，这并不符合物理手部结构。因此通过建立良好的手部模型，开发手部关节特性，可以降低手部模型重合等现象。

综上所述，本章阐述了手部构型与姿态表示方法，通过对比不同姿态表示方式，建立基于四元数的姿态表述模型，该模型能够支持多种典型手部构型。此外，本文分析了可穿戴设备的误差原因，通过结合手部关节约束，提出了一种全手姿态优化方法，该方法能使结果符合手部物理结构，降低了可穿戴设备的错位误差，并通过实验证实了该方法的有效性。

### 3.1 姿态表示与关节特性

本文将姿态解释为物体在空间中的朝向，不包含物体在空间中的位置信息。姿态可以用多种方法来描述，例如欧拉角、旋转矩阵、旋转向量、四元数等。手部姿态可以用手掌与骨节空间朝向表示，因此可以将手部模型中骨节看作空间中的刚体，关节将手部骨节结合到一起，构成手部模型。在手部姿态求解时，可以通过将传感器附着于骨节上进行姿态求解，但简单的组合无法表达精准的手部姿态，需要挖掘关节特性。

本小节分析了姿态描述方法，研究了一种以四元数为基础的手部模型。其次描述了手部关节的运动特性与生理约束，为手部跟踪模型的优化提供基础。

#### 3.1.1 姿态描述方法

空间中刚体的姿态可以描述为物体的空间朝向，可以用从初始坐标系旋转到物体所在坐标系的旋转变换过程表示。刚体的姿态表述方法有许多种<sup>[60]</sup>，各自存在其优缺点，需要根据使用场景的需求，选择合适的描述方式建立模型。在本文中，人体手部可以视为一个复杂的关节链系统，由各个骨节组成，骨节与骨节之间通过关节连接。手部的骨节在交互过程中不会发生形变，可以视为空间中的刚体，因此对于刚体的姿态描述也可以用作手部姿态描述。

刚体的姿态通常可以使用欧拉角、旋转矩阵或四元数来表示。

欧拉角是一种用于描述三维空间中物体姿态的方法，它通过三个旋转角度（通常是滚动、俯仰和偏航）来表示姿态，每个角度对应于物体绕固定坐标轴的旋转，这种表述方式方便刚体姿态的可视化。

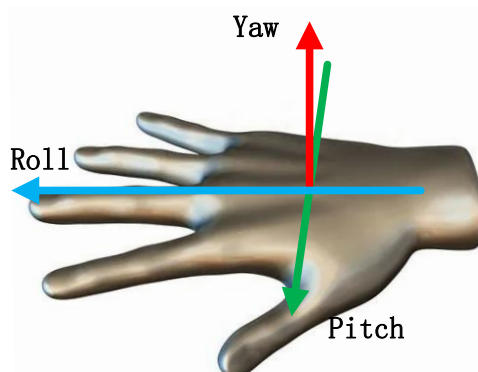


图 3-1 欧拉角旋转轴示意图

如图 3-1 所示，欧拉角表达直观简洁，仅需要三个角参数就可以确定刚体在三维空间的方向，这符合人们根据日常经验对姿态变化的感知习惯。不同欧拉角定义方式相对独立，每一个角度变化都可以理解为一个基本的坐标轴旋转，这与许多实际机械运动形态类似。例如飞机的驾驶操作即对应欧拉角中的滚转、俯仰和偏航三个变化。欧拉角运算简单，在手工计算和工程实现中都比较方便。基于这些优点，欧拉角在许多工程领域得到大范围运用。使用欧拉角描述刚体姿态也存在一些缺点，例如万向锁问题，即当两个旋转轴趋于平行时，会导致失去一个自由度，使得刚体在特定方向上无法自由旋转。此外使用欧拉角描述刚体姿态时，不同旋转轴顺序可能会导致不同的空间朝向，在使用时需要设定旋转轴顺序。

旋转矩阵也是描述三维空间刚体姿态的重要工具。旋转矩阵描述刚体姿态的原理基于线性代数和几何变换概念。通过一个  $3 \times 3$  的矩阵表示旋转变换，该矩阵具有正交性的特点，确保旋转前后不会改变物体形状等属性。刚体的姿态变换可通过旋转矩阵将初始坐标系下的点变换到新的坐标系。多次旋转则可以用旋转矩阵的乘积来表示最终的刚体姿态。

使用旋转矩阵描述刚体姿态的优点在于其简洁而有效地捕捉了刚体在三维空间中的旋转变换，通过旋转矩阵的乘积，可以方便地表示连续的旋转变换，适用于复杂的运动场景。在机器人学中，姿态控制和运动规划也经常使用旋转矩阵来表示末端执行器的姿态。另外，飞行器的导航系统也常利用旋转矩阵描述飞行器的方向变化。使用旋转矩阵描述刚体姿态的缺点，由于旋转矩阵的维度为  $3 \times 3$ ，当进行复杂运动时，其表达式相对复杂，并且表示同一刚体姿态时对应旋转矩阵不是唯一的，存在多个等价的表示方式，这可能导致一些混淆和不一致性。

除欧拉角与旋转矩阵外，研究者还将四元数用于姿态描述。19 世纪，爱尔兰数学家威廉哈密顿提出了四元数的概念，四元数作为复数的推广，由一个实部和三个虚部组成，

三个虚部之间正交。描述三维空间中两个坐标系的相对姿态时，姿态四元数 $\mathbf{q}$ 可以用公式 3-1 表示。

$$\mathbf{q} = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} \cos(\frac{\theta}{2}) \\ e_x \sin(\frac{\theta}{2}) \\ e_y \sin(\frac{\theta}{2}) \\ e_z \sin(\frac{\theta}{2}) \end{bmatrix} \quad (3-1)$$

在式 3-1 中， $\theta$ 为相对旋转角度， $e_x$ 、 $e_y$ 、 $e_z$ 为单位旋转轴的各个分量。在执行连续的旋转操作时，例如单位四元数 ${}^A_B \mathbf{q}$ 表示从 $B$ 坐标系旋转到 $A$ 坐标系的旋转过程，可以通过借助中间状态 $G$ 坐标系完成，先将 $B$ 坐标系通过旋转转化为 $G$ 坐标系，再通过 $G$ 坐标系旋转到 $A$ 坐标系，四元数乘法很好的实现了这一过程，见公式 3-2。

$${}^A_B \mathbf{q} = {}^A_G \mathbf{q} \otimes {}^G_B \mathbf{q} \quad (3-2)$$

在公式 3-2 中， ${}^A_G \mathbf{q}$ 为 $G$ 坐标系旋转到 $A$ 坐标系的四元数表示， ${}^G_B \mathbf{q}$ 为 $B$ 坐标系旋转到 $G$ 坐标系的四元数表示， $\otimes$ 为四元数乘法标识符。在研究三维空间中的刚体运动时，不仅需要考虑刚体的静止姿态，还需要描述随时间发生的姿态变化和旋转。为了达到这一目的，必须建立刚体姿态的微分方程模型。刚体的姿态变化可被视为刚体所在坐标系相对于空间坐标系的转动，而这种转动的速度即为刚体的角速度。因此，刚体的姿态变化方程与角速度之间存在直观而密切的对应关系。采用四元数表达刚体姿态时，根据刚体角速度构造四元数微分方程的一般形式如公式 3-3。

$$\dot{\mathbf{q}} = \frac{1}{2} \mathbf{q} \otimes \boldsymbol{\omega} = \begin{bmatrix} -\frac{1}{2}(\omega_x q_1 + \omega_y q_2 + \omega_z q_3) \\ \frac{1}{2}(\omega_x q_0 + \omega_y q_3 - \omega_z q_2) \\ \frac{1}{2}(\omega_y q_0 + \omega_z q_1 - \omega_x q_3) \\ \frac{1}{2}(\omega_z q_0 + \omega_x q_2 - \omega_y q_1) \end{bmatrix} \quad (3-3)$$

在公式 3-3 中， $\dot{\mathbf{q}}$ 为单位四元数，表示刚体瞬时姿态， $\boldsymbol{\omega}$ 为刚体坐标系下的角速度三维向量， $\boldsymbol{\omega} = [0, \omega_x, \omega_y, \omega_z]$ ， $\otimes$ 表示四元数乘法。上述四元数微分方程直观地反映了刚体运动学之间的内在联系。通过求解这一微分方程组，即可对刚体的运动轨迹进行解析，也为基于四元数的姿态估计和控制提供了理论基础。

四元数最初被应用到计算机图形学中，用于实现三维场景中的旋转变换与插值。随着研究的深入，人们发现四元数在机器人学、导航、控制等领域也有广泛的应用价值。例如在无人机的姿态控制中，四元数可以避免欧拉角的万向锁问题，实现平稳的姿态跟踪。四元数为刚体姿态建模提供了一种有效的数学工具，它与欧拉角、旋转矩阵等传统方法可以互相转换，也可以根据应用场景需要进行替代。

四元数方法能够获得线性的姿态运动学方程，直观地表达三维空间旋转并便于进行旋转过程插值，利于手部运动轨迹的估计。结合三种描述方式的适用场景与特点，本文使用四元数作为手部姿态的描述形式。

### 3.1.2 手部关节特性

本小节介绍了手部自由度与常见手部构型，研究了一种于四元数的三指简化手部模型，用于后续手势跟踪与手势交互。本小节虽然选用三指简化构型，但本文提出方法不局限于三指的姿态估计，可以拓展并支持全手姿态估计。人体手部由骨节姿态和关节姿态共同组成，为了更好的描述手部姿态，根据需求选择并创建合适的手部模型。手部结构包含许多小骨和关节，具有不同的自由度。但是从运动协调的角度来看，手部运动可化简为各个关节与骨节的组合运动，这使得可以建立一个相对简化的手部模型，同时保留手部运动的主要特征。

手部关节包括各个指掌关节和指间关节，除拇指外的其他手指关节又包括指间近端关节和指间远端关节两个关节，而且在不同的研究中，手部运动模型也不一致。有研究者提出了一种比较精确的具有 31 个自由度手部运动模型<sup>[61]</sup>，其中每个手指都具有 5 个自由度，手的位置和方向 6 个自由度。也有研究人员将手部划为 24 个自由度的运动模型<sup>[62]</sup>，自由度包括拇指 5 个自由度，余下手指各 4 个自由度，手的姿态 3 个自由度。这些复杂的关节结构与运动学链协同工作，使手部可完成精细而复杂的动作。通过组合手掌的旋转角度以及每个手指关节的弯曲角度，可以精确描述复杂的手部姿态，这为分析复杂的手势提供了基础。研究者 Wen<sup>[37]</sup>等人根据手指在日常生活中的功能权重占比，分析食指拇指和中指在整手功能占比为 72%，构造了一种三指简化模型。

具体来说，人手包含一个球形手腕的结构，实现手腕三轴运动，然后是五个手指，每个手指由多个关节串联而成，手部关节与指骨示意图如图 3-2 所示。

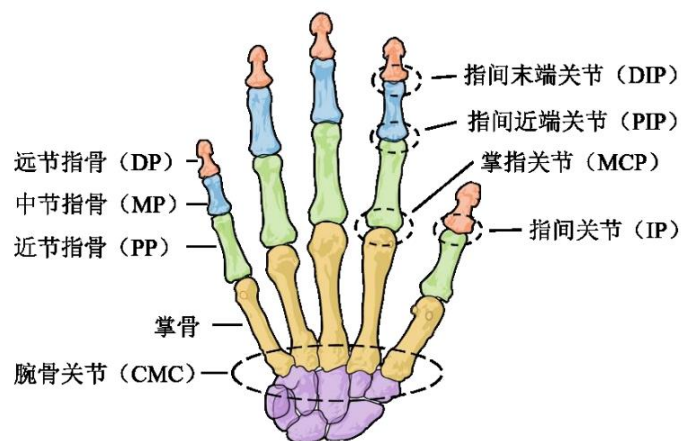


图 3-2 手部关节与指骨示意图<sup>[63]</sup>

尽管每个关节都具有一定自由度，但在许多抓握模式中，多个关节会协同工作，形成协同运动模式，因此可以将每个手指简化为 2 到 3 节刚性物体链，每一节相当于一个

刚性物体,这种模型化表示降低了模型复杂度。基于以上思路,研究者提出了许多手部运动学模型,这些模型在保留手部主要运动特征的同时,具有少量参数和简单的运动学方程。这类模型与复杂手部模型相比,在大范围手部姿态估计算法中表现相近,因此,设计合适的简化手部模型,对空间手部姿态跟踪与预测问题是至关重要的。本文结合交互需求,选用三指构型作为全手行为跟踪的手部结构模型,由拇指、食指和中指组成。

手部结构不仅仅是骨节的互相连接,人体手部还有复杂的肌肉与神经相互作用,形成手部的关联运动。其中手部关节表现出具有活动范围的特点,超出关节活动范围会造成手部损伤。Cobos 等人<sup>[64]</sup>研究了人手各关节的静态运动范围约束,为建立符合生物学特性的手部姿态估计模型奠定了解剖学基础。具体来说,他们通过实验测量了受试者手部在无外力作用下的最大自主运动范围。这种约束反映了人手的骨骼关节、肌肉、肌腱等组织结构的生物力学特性。

表 3-1 手部关节活动范围<sup>[64]</sup>

手指	关节	弯曲角度	外展角度
拇指	MCP	[0°~75°]	[0°~5°]
	IP	[-5°~75°]	0°
食指	MCP	[-30°~90°]	[0°~60°]
	PIP	[0°~110°]	0°
	DIP	[-5°~90°]	0°
中指	MCP	[-30°~90°]	[0°~45°]
	PIP	[0°~110°]	0°
	DIP	[-5°~90°]	0°
无名指	MCP	[-30°~90°]	[0°~45°]
	PIP	[0°~120°]	0°
	DIP	[-5°~90°]	0°
小指	MCP	[-30°~90°]	[0°~50°]
	PIP	[0°~135°]	0°
	DIP	[-5°~90°]	0°

通过表 3-1 可以看出,手部包含多个关节,其中可以划分为两类:(1)指间关节(IP, Interphalangeal Joint)包括近端指间关节(PIP, Proximal Interphalangeal Joint)和远端指间关节(DIP, Distal Interphalangeal Joint);(2)掌指关节(MCP, Metacarpophalangeal Joint)。指间关节被归类为单自由度关节,其运动仅限于围绕一个轴进行旋转。相对而言,其运动模式较为简单。与之不同的是,掌指关节被定义为多自由度关节,能够实现较为复杂的运动。此外,掌指关节之间受到覆盖手部的肌肉的制约,不同方向上的运动不能被视为完全独立。肌肉作用导致多个掌指关节发生协同运动,相互之间存在共同的约束。相对于指间关节,掌指关节的运动模式更为复杂。手部约束与人体生物力学特性相关,每个人会有一定的个体差异。但总体上主动约束反映了人类生物结构对手部运动的协调规律。

在设计手部姿态估计算法时,需要考虑不同类型的约束,建立能够反映真实手部运动协同规律的模型。加入约束知识,可以提高姿态估计的准确性,也使得预测系统能够生成更符合生物学特性的手部运动。

综合利用手部约束信息构建手部运动学模型,恰当利用先验的生物学知识,可以使得全手行为跟踪算法在特定应用场景中获得更好的效果,在特定场景中,这些基本约束为姿态估计系统提供了重要的先验知识。例如佩戴数据手套时由于传感器与骨节之间必定存在间隙,附着传感器与骨节形成方向错位而导致姿态估计不准确,融合约束知识会矫正误差。同时,深入理解由人体解剖结构导致的静态运动范围限制对于构建仿生手模型具有重要价值。虽然这些约束不一定普遍提高姿态估计的精度,但它们提供了有用的先验知识,以更准确地模拟生物运动。这种生物约束的引入也为人机交互、运动控制等相关领域提供了宝贵的启发。

### 3.2 全手行为跟踪方法

在使用数据手套进行手部行为跟踪时,可将手套上的传感器姿态视为相应骨节的姿态。因此,通过传感器姿态求解算法并结合手部模型,可以实现全手行为跟踪。本节介绍了一种 IMU 姿态估计算法,基于此算法将手掌和手指骨节的姿态作为系统状态,构建了线性卡尔曼滤波模型。在该模型中,利用角速度更新系统状态,同时使用磁信号与加速度拟合的姿态作为观测值,实现进行全手行为跟踪。

#### 3.2.1 IMU 姿态求解方法

本小节介绍了一种融合磁力计、陀螺仪、加速度计的 IMU 姿态求解方法,为全手行为跟踪方法提供一种理论支持。在进行姿态估计时,单一类型的传感器往往会因为自身误差的累积而产生明显的漂移和抖动现象。这主要有两个主要原因。首先,传感器本身的制造工艺误差是一个重要因素,例如在惯性传感器存在零偏误差和初始离散误差<sup>[37]</sup>。其次,传感器在运行环境中所受到的干扰也是一个重要考虑因素。例如,温度变化可能影响陀螺仪的均值,而磁场干扰则可能影响磁强计的准确性。这些环境因素同样会导致单一传感器的输出存在明显的漂移。因此,针对这些问题,采用多传感器融合的方法可以有效减小误差,提高姿态估计的准确性和稳定性。

研究者在解决手部姿态问题时广泛采用传感器融合策略。其中一种典型的方法是在滤波框架下对数据进行优化融合,该模式综合两种传感器各自的优势,以获得更准确且稳定的手部姿态输出。设计高性能的多传感器融合方案被认为是提升手部姿态估计准确性的重要途径。传感器的测量值是基于传感器的姿态给出的,即可以通过数学模型推导出初始状态到传感器姿态的姿态差异,从大地坐标系到传感器坐标系的用四元数表示的旋转过程。本方法通过将姿态推算分为重力与传感器两个坐标系系统,并为了方便标记传感器坐标系与大地坐标系,称为 $L$ 和 $G$ 。

在传感器静止的状态下,加速度计只存在重力加速度,在本文中用 ${}^L\boldsymbol{a}$ 表示,而大地

坐标系的重力加速度是一个常量，且用向量 ${}^G\mathbf{g}$ 表示，满足只在竖直方向上有数值，而其他分量没有加速度存在，为了计算的规范性，将加速度计的测量值与重力加速度都进行标准化处理。同样，磁力计可以测量出传感器所在的磁场数据，本文用 ${}^L\mathbf{m}$ 表示传感器坐标系下的单位向量，用 ${}^G\mathbf{h}$ 表示真实的地球磁场中的单位向量，即 Z 轴竖直向上，并且 X 轴指向正北方向。

单位四元数可用于操作三维向量的旋转，本文用四元数 ${}^L_G\mathbf{q}$ 来表示大地坐标系 $G$ 中向量与传感器坐标系 $L$ 的单位向量的旋转。所以 ${}^L_G\mathbf{q}$ 应满足将传感器坐标系下的单位向量 ${}^L\mathbf{a}$ 转化为大地坐标系下的单位向量 ${}^G\mathbf{g}$ ，将传感器坐标系下的单位向量 ${}^L\mathbf{m}$ 转化为大地坐标系下的单位向量 ${}^G\mathbf{h}$ 。故写成代数方程的形式如下式 3-4。

$$\begin{cases} \mathbf{R}^T({}^L_G\mathbf{q}){}^L\mathbf{a} = {}^G\mathbf{g} \\ \mathbf{R}^T({}^L_G\mathbf{q}){}^L\mathbf{m} = {}^G\mathbf{h} \end{cases} \quad (3-4)$$

通过式 3-4 可以确定 ${}^L_G\mathbf{q}$ 的代数解，即获得 ${}^L\mathbf{a}$ 与 ${}^L\mathbf{m}$ 的函数。通过加速度获得向量轴介于地球坐标系 $G$ 与传感器坐标系 $L$ 的旋转四元数 ${}^L_G\mathbf{q}_{acc}$ ，通过磁场信息可以得到其余两轴介于地球坐标系 $G$ 与传感器坐标系 $L$ 的旋转四元数 ${}^L_G\mathbf{q}_{mag}$ ，为了方便计算，引入了一个中间过程坐标系 $I$ 以连接两式，故中间过程坐标系下的向量表示通过 ${}^L_G\mathbf{q}_{acc}$ 矫正后得到四元数具有无偏差的 Z 轴，然后再用 ${}^I_G\mathbf{q}_{mag}$ 对中间过程坐标系的向量的余下两轴进行校准，故 ${}^L_G\mathbf{q}$ 满足式 3-5。

$${}^L_G\mathbf{q} = {}^L_I\mathbf{q}_{acc} \otimes {}^I_G\mathbf{q}_{mag} \quad (3-5)$$

如此便建立了对于四元数的求解理论部分，通过重力场进行竖直方向的矫正，求解出 ${}^L_I\mathbf{q}_{acc}$ 的代数解，解算出中间坐标系下的向量，然后利用磁场信息校准其余两轴的方向，求解出 ${}^I_G\mathbf{q}_{mag}$ 的代数解。接下来讲述具体的计算过程。重力场测定竖直方向并求得 ${}^L_I\mathbf{q}_{acc}$ 代数解。因为磁力计较容易收到周围环境的干扰，加速度计容易受到运动的干扰，因此结合磁力计与加速度计进行姿态求解的方法有较高的噪声，该方法得到一个有较大的噪声的传感器姿态。

基于陀螺仪的姿态求解则是通过三轴陀螺仪采集角速度矢量，由查阅文献知，刚体的运动服从角速度矢量微分方程式 3-6。

$$\mathbf{q}_{t+1} = \Phi(\boldsymbol{\omega}, \Delta t)\mathbf{q}_t \quad (3-6)$$

在式 3-6 中， $\mathbf{q}_{t+1}$ 与 $\mathbf{q}_t$ 分别表示在  $t+1$  时刻和  $t$  时刻的物体姿态。 $\Phi(\boldsymbol{\omega}, \Delta t)$ 则表示对角速度 $\boldsymbol{\omega}$ 和采样周期 $\Delta t$ 进行积分后的状态转移矩阵。通过陀螺仪采集到的连续的数据，即可推导出连续的姿态四元数。但是由于陀螺仪本身有测量误差，并且测量误差随着时间的增长，积分后的误差累计越来越大。在短时间内是稳定连续的，但是长时间的积分会导致整体误差比较大。

因此有许多研究者进行姿态的计算。互补滤波算法利用陀螺仪、加速度计以及磁力计所具有的互补性能进行多传感器数据融合。该算法通过建立反馈控制器，将姿态解算过程中的误差反馈进行补偿，以校正陀螺仪积分漂移引起的低频误差。此外，该算法还

保留了陀螺仪的高带宽性能，能够对高动态运动进行准确捕捉。通过这种方式，互补滤波算法兼顾了低频和高频信号，满足了精确姿态解算的动态和静态需求。卡尔曼滤波器基于贝叶斯递归估计框架<sup>[65]</sup>，是一个连续时间线性系统的最优递归数据处理算法。它利用目标系统的状态方程建立动力学模型，并构建过程噪声和观测噪声的协方差矩阵。通过预测和更新递归步骤，卡尔曼滤波器对系统状态进行估计，并通过反馈调节获得最优解。相较于互补滤波，卡尔曼滤波器充分利用所有传感器信息，通过统计优化获得最佳状态估计。此外，其数学框架的严格性也使滤波过程更加稳定可靠。总体来说，这些多传感器融合算法各有优势，应根据实际需求进行选用与设计。

### 3.2.2 全手行为跟踪方法

本小节基于3.2.1节中描述的传感器姿态求解方法，研究了一种全手行为跟踪方法。根据手部姿态的定义和特性构建卡尔曼滤波系统，该系统其包含系统状态方程与观测方程两部分，观测系统是根据加速度计和磁力计的数据对手部各骨节姿态进行解算，而系统状态方程则是根据通过陀螺仪采集到的连续的数据，进行二次积分推导骨节姿态。在本方法中，所提到的骨节姿态均基于贴合在骨节上的惯性传感器，得到的全手姿态亦为附着的传感器的姿态，本方法假定传感器姿态即为附着的骨节姿态。

在本文的手部姿态估计方法中，将8个手部骨节姿态信息作为系统状态，对三指关节进行约束空间求解，每个骨节姿态以四元数形式表示，系统状态记作 $\mathbf{x}$ ，如公式3-7所示。

$$\mathbf{x} = [H \quad T_{FM} \quad T_{PP} \quad T_{DP} \quad F_{PP} \quad F_{MP} \quad M_{PP} \quad M_{MP}]^T \quad (3-7)$$

在公式3-7中， $H$ 表示手的姿态，贴合食指与中指的掌骨， $T$ 、 $F$ 、 $M$ 分别表示手指为拇指、食指和中指， $FM$ 、 $PP$ 、 $MP$ 、 $DP$ 则分别表示拇指的掌骨、近端指骨、中端指骨、远端指骨。

通过相邻指骨姿态的数值运算，可以得到相邻指骨间关节的姿态，对关节姿态进行约束空间下求解，获取系统状态的先验估计与获取观测系统的值。该系统包含及系统方程和观测方程两部分。系统方程使用陀螺仪进行姿态的变化求解，而观测系统则没有累计误差、利用加速度计与磁力计求解得到的四元数姿态。故给出系统的先验估计值和观测系统得到的观测值与系统状态的关系表达式3-8。

$$\mathbf{x}_t^- = A\mathbf{x}_{t-1} + \mathbf{w}_t \quad (3-8)$$

在式3-8中， $\mathbf{x}_{t-1}$ 表示 $t-1$ 时刻的系统的当前姿态， $\mathbf{x}_t^-$ 表示将 $t$ 时刻的系统状态作为先验知识，通过角速度的积分得到系统状态得先验估计。 $A$ 表示通过角速度计算得到的状态转移矩阵，将 $t-1$ 时刻的状态转移到 $t$ 时刻。 $\mathbf{w}_t$ 则由传感器的陀螺仪噪声组成。

观测方程则不随时间增加而累计误差，根据加速度计和磁力计的数据对传感器姿态进行解算。本文基于加速度计和磁力计的姿态解算通过两次旋转完成，第一次旋转是通过加速度计的输出解算得到，相应的四元数描述了旋转过程中的翻滚角和俯仰角信息。



第二次旋转是通过磁力计的输出解算得到，相应的四元数描述仅包含了传感器的偏航角，第二次旋转基于第一次旋转之后的坐标系。在得到观测系统的旋转姿态之后需要与当前系统状态进行转换，观测系统与当前系统状态的状态转化表达式如式 3-9 所示。

$$\mathbf{z}_{k+1} = H\mathbf{x}_{k+1} + \mathbf{v}_{k+1} \quad (3-9)$$

在式 3-9 中， $H$  表示系统状态姿态坐标系与观测系统姿态坐标系之间的转换矩阵，在本文中为单位阵， $\mathbf{z}_{k+1}$  表示  $k+1$  时刻下观测系统所得的姿态值，与系统状态组成相同。观测系统首先通过加速度计计算了传感器坐标系和大地坐标系下向量之间的差异。随后，利用磁力计获取了传感器坐标系下的向量，并使其方向与大地坐标系下的向量一致。接着，通过解算方程组计算了两种坐标系下偏航角的差异。结合两个四元数，求解传感器坐标系与大地坐标系之间的旋转过程，表示传感器坐标系与大地坐标系的姿态差异，也即为所贴合指骨的姿态。

在得到系统观测姿态值与系统先验姿态值后，分别计算其对应的误差协方差矩阵<sup>[65]</sup>，系统先验误差协方差增益可以将陀螺仪误差转换得到，观测系统的误差协方差矩阵则根据磁力计与加速度计的误差通过泰勒的一阶展开式给出，根据其对应的误差协方差矩阵，计算并更新卡尔曼增益  $K$ ，进而将观测系统计算所得手部姿态与系统先验手部姿态融合求解，得到最优手部姿态值。

$$\mathbf{x}_{k+1} = \mathbf{x}_{k+1}^- + K(\mathbf{z}_{k+1} - \mathbf{x}_{k+1}^-) \quad (3-10)$$

在式 3-10 中， $\mathbf{x}_{k+1}$  表示当前系统  $k+1$  时刻下求解所得最优手部姿态值， $\mathbf{x}_{k+1}^-$  表示在  $k+1$  时刻下手部姿态的先验估计值， $K$  表示通过误差协方差矩阵获得的卡尔曼增益， $\mathbf{z}_{k+1}$  表示  $k+1$  时刻下观测系统所得的姿态值。在利用卡尔曼增益完成系统状态值的更新之后，进一步对系统状态估计的误差协方差进行更新，达到动态校正系统状态值的目的。

### 3.3 全手跟踪优化方法

本小节首先分析了全手行为跟踪的误差，结合人体手部特性研究了一种融合人手生理约束的全手跟踪优化方法，全手行为跟踪优化算法结构如图 3-3 所示。

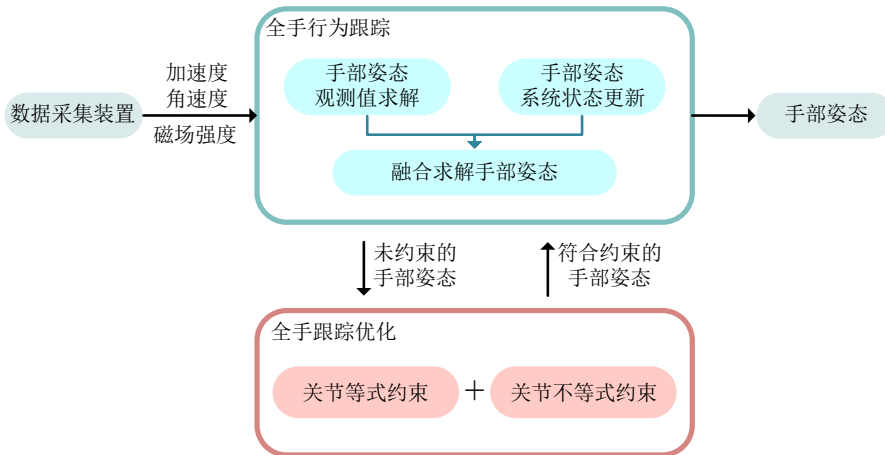


图 3-3 全手行为跟踪优化算法结构图

手部姿态估计经常需要考虑生理学上的手部关节约束，大量解剖学实验表明，手部运动往往受骨骼、肌肉和软组织的生理结构限制。而将这些限制关系纳入全手行为跟踪的方法中，则可以得到符合人体构造的手部姿态。本研究所用全手跟踪优化方法基于这一思路，使用数据手套来捕捉手部运动，并结合 3.1 节阐述的生理学约束关系进行数据融合，以提高姿态估计的准确性。对手势交互中存在的误差进行误差分析，本节将介绍交互过程中全手跟踪的误差，并引入关节约束特性对 3.2 小节中的行为跟踪方法进行优化，从而降低运动过程中的误差。

### 3.3.1 全手行为跟踪的误差分析

全手行为跟踪中的误差可以分为系统误差、环境误差、运动过程误差等。姿态估计系统的误差，传感器和姿态估计系统本身具有一部分误差，由估计方法的误差和传感器误差组成，可以通过加入零均值高斯白噪声和融合偏差的方式降低这一部分误差。而环境误差通常是由空间中磁场不均匀导致的，环境磁场不均匀影响磁力计的读数，任何磁场变化都会对姿态估计产生影响，通常会通过事先测量环境磁场分布，建立磁场映射校正磁力计测量，降低该部分环境误差。

可穿戴设备由于其可拆卸的特性，其固定方式无法完全贴合手部，因此在运动过程中难免会发生相对位移。运动过程误差是在手部运动时产生的，其表现受到不同佩戴方式、手套材质以及传感器分布的影响。与人体皮肤和肌肉富有弹性和韧性不同，数据手套的材质通常缺乏这些特性，导致在进行手部动作时受到材质延展性的限制。当骨节驱动传感器运动时，传感器受到手套的约束，从而引起与真实手部姿态的误差，尤其在执行大范围的移动时，这种误差变得更为明显。数据手套利用惯性传感器与骨节绑定以记录手指运动，为手部姿态估计提供关键数据。然而，惯性测量单元通常固定在手指骨节表面，手指活动时手套与皮肤之间发生相对滑动，皮肤与骨骼之间也有相对滑动，从而使得 IMU 的位置相对于骨骼发生变化。这种相对错位现象在手指进行大幅度弯曲运动时尤为明显，佩戴错位的示意图如图 3-4 所示。

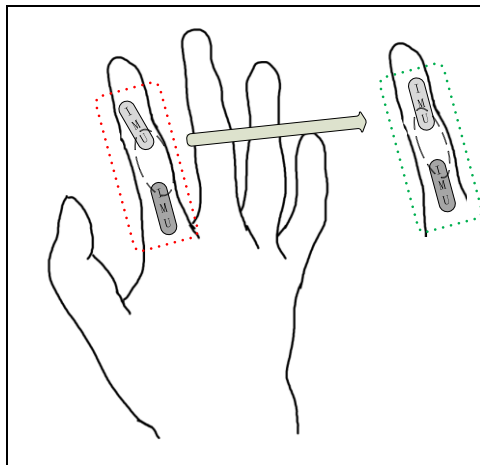


图 3-4 可穿戴设备佩戴错位示意图

在图 3-4 中，左侧手部为佩戴错位图，右侧手指为正确固定时示意图。在手势交互过程中，手指之间的直接接触会对姿态估计产生一定干扰，具体来说，当用户同时弯曲多个相邻手指时，这些手指的末端可能发生相互接触。因手套表面绑定有测量各手指关节弯曲角度的弯曲传感器，手指之间的直接接触会对相应的弯曲传感器产生额外的机械力，导致关节角度测量结果出现错误。例如，当食指和中指同时并拢弯曲时，两指的接触会相互施加额外力，使惯性传感器测量的中指和食指的弯曲角度比实际弯曲角度要小。这种手指直接接触引起的弯曲角度测量误差会导致最终手指姿态跟踪的误差。此外，手指之间的密集接触还可能导致手套表面的惯性传感器发生错位，这也会对后续的姿态求解产生较大误差。

在运动测量中，惯性传感器的佩戴位置误差是影响测量精度的一个关键因素。这主要是由于人体手部关节运动的复杂性所引起的。举例来说，在进行手部运动时，指尖经常会伴随手指的侧转和扭转而移动，导致佩戴在手指上的惯性传感器位置发生偏移。这种运动过程中传感器与骨节之间的错位问题会影响对手部姿态的准确估计。在上述过程中，如果佩戴方式没有得到有效纠正，这类误差将在用户执行手势过程中持续存在。为解决这类问题，本研究将关节的特性引入手部跟踪系统中，以降低运动过程误差。

### 3.3.2 构建手部非线性约束模型

为更好地将手部约束纳入系统模型，考虑手部约束的特点和融合方法需求，本小节将第 3.2 节中的手部关节角度约束转换为基于四元数姿态的等式和不等式约束，并将其融入手部跟踪模型中。在本文中，手部关节被分为单自由度关节和多自由度关节。其中，单自由度关节包括拇指 IP 关节以及其他手指的 DIP 和 PIP 关节，这类关节仅能进行弯曲动作，具有一个自由度。而多自由度关节则包括各个手指的 MCP 关节，这类关节不仅能进行弯曲动作，还能进行外展/内收的伸展动作。本节分别针对单自由度关节和多自由度关节进行建模，并将关节执行动作角度的范围约束转化为作用于系统状态的数值约束，再选用求解方法进行求解。

本文将手部约束作为先验的知识融合到卡尔曼滤波系统的状态方程中，求解得到符合手部约束先验姿态值。卡尔曼滤波系统有系统状态方程与观测方程两部分，观测系统下的手部姿态是根据加速度计和磁力计的数据对传感器姿态进行解算，而系统状态方程则是根据通过陀螺仪采集到的连续的数据，进行二次积分推导姿态四元数，测量误差随着时间的增长，积分后的误差逐步增大，在本文中，结合手部约束可以减少这种误差，融合关节约束的系统状态方程可以描述为如式 3-11 所示。

$$\mathbf{x}_{k+1}^- = A\tilde{\mathbf{x}}_k + \mathbf{w} \quad (3-11)$$

在式 3-11 中， $\mathbf{x}_{k+1}^-$  表示在  $k+1$  时刻下手部姿态的先验估计值， $A$  则表示由角速度积分得到的状态转移矩阵， $\mathbf{w}$  表示符合高斯分布的系统过程噪声， $\tilde{\mathbf{x}}_k$  则表示在  $k$  时刻下满足手部约束的手部姿态值，也是上一时刻的符合生理学约束的系统状态。

由于本文的系统状态为手部骨节姿态，因此需要通过相邻手部骨节姿态来获取该关节的姿态，假设手部上的两个相邻指骨标记为 $L1$ 和 $L2$ ，两指骨所形成的关节 $Joint$ 则需要通过两骨节运算表示。

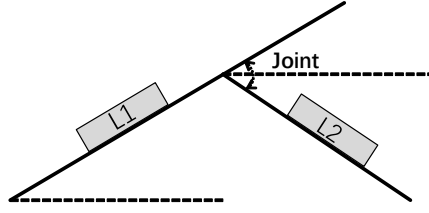


图 3-5 通过相邻骨节求解关节姿态示意图

在图 3-5 中， $Joint$  关节姿态可以由骨节  $L1$  基于骨节  $L2$  的旋转过程来表示，即为将骨节  $L1$  坐标系转换为骨节  $L2$  坐标系的旋转过程，称为该关节的姿态，两者的姿态都是通过大地坐标系所表示，例如  $L1$  的姿态则是从大地坐标系旋转到传感器坐标系  $L1$  的旋转过程，所以将大地坐标系作为桥接坐标系，在此状态下再进行大地坐标系到传感器坐标系  $L2$  的旋转，即得到从传感器坐标系  $L1$  旋转到传感器坐标系  $L2$  的旋转过程，也即为骨节  $L1$  基于骨节  $L2$  之间所形成的关节，如公式 3-12 所示。

$$\mathbf{q}_{Joint} = \mathbf{q}_{L2} \otimes (\mathbf{q}_{L1})^{-1} = \begin{bmatrix} h_1(\mathbf{q}_{L1}, \mathbf{q}_{L2}) \\ h_2(\mathbf{q}_{L1}, \mathbf{q}_{L2}) \\ h_3(\mathbf{q}_{L1}, \mathbf{q}_{L2}) \\ h_4(\mathbf{q}_{L1}, \mathbf{q}_{L2}) \end{bmatrix} \quad (3-12)$$

在公式 3-12 中， $h_i$  代表姿态四元数系数的函数， $i$  对应该关节姿态四元数的系数编号， $\mathbf{q}_{L1}$  表示  $L1$  骨节所对应的姿态。上述过程得到系统估计所得的关节姿态，要对关节姿态进行约束，则需要得到每个关节姿态的约束范围，也即将针对关节的角度约束，转化为针对关节的姿态约束。

本文将关节约束分为了单自由度关节约束与多自由度关节约束，单自由度关节约束主要特征就是只存在一个方向上的转动，在单自由度关节约束建模中，假定该关节的弯曲角度为  $\theta$ ，对应姿态绕  $x$  轴进行旋转，则该关节的姿态如公式 3-13 所示。

$$\mathbf{q}_{Joint} = \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) \\ 0 \\ 0 \end{bmatrix} \quad (3-13)$$

通过公式 3-13 可以得到弯曲角度为  $\theta$  的四元数姿态，假定该角度的取值范围为  $[\theta_0, \theta_1]$ ，则通过简单求导即可得到各个值的最大值与最小值范围，结合公式 6 可以即可得到姿态的范围约束结果，如公式 3-14 所示。

$$\mathbf{q}_{joint} = \begin{bmatrix} \min(\cos(\frac{\theta}{2})) \leq h_1(\mathbf{q}_{L1}, \mathbf{q}_{L2}) \leq \max(\cos(\frac{\theta}{2})) \\ \min(\sin(\frac{\theta}{2})) \leq h_2(\mathbf{q}_{L1}, \mathbf{q}_{L2}) \leq \max(\sin(\frac{\theta}{2})) \\ h_3(\mathbf{q}_{L1}, \mathbf{q}_{L2}) = 0 \\ h_4(\mathbf{q}_{L1}, \mathbf{q}_{L2}) = 0 \end{bmatrix} \quad (3-14)$$

在公式 3-14 中,  $h_1(\mathbf{q}_{L1}, \mathbf{q}_{L2})$ 和 $h_2(\mathbf{q}_{L1}, \mathbf{q}_{L2})$ 为具有最大值与最小值的不等式约束,  $h_3(\mathbf{q}_{L1}, \mathbf{q}_{L2})$ 和 $h_4(\mathbf{q}_{L1}, \mathbf{q}_{L2})$ 则为固定值的等式约束, 因此可以将这个拆分成等式约束和不等式约束两部分。结合 3.2 小结内容, 通过将表 3-1 中的各个角度约束范围代入公式 3-14 中, 得到针对关节四元数的值约束, 如表 3-2 所示。

表 3-2 手部关约束转化为数值约束的一般形式

手指	关节	关节姿态约束表达式
拇指	MCP	$0 \leq q_1, 0 \leq q_2 < 0.6088,$ $-0.0133 < q_3 < 0.0133, -0.0218 \leq q_4 < 0.0218$
	IP	$0.7933 \leq q_1 \leq 1, -0.0436 < q_2 < 0.6088, q_3 = q_4 = 0$
食指	MCP	$0 \leq q_1, -0.2588 < q_2 < 0.7071,$ $-0.183 < q_3 < 0.183, -0.2588 \leq q_4 < 0.2588$
	PIP	$0 \leq q_1, 0 \leq q_2 \leq 0.8192, q_3 = q_4 = 0$
	DIP	$0 \leq q_1, -0.0436 < q_2 < 0.7071, q_3 = q_4 = 0$
中指	MCP	$0 \leq q_1, -0.2588 < q_2 < 0.7071,$ $-0.138 < q_3 < 0.138, -0.1951 \leq q_4 < 0.1951$
	PIP	$0 \leq q_1, 0 \leq q_2 \leq 0.8192, q_3 = q_4 = 0$
	DIP	$0 \leq q_1, -0.0436 < q_2 < 0.7071, q_3 = q_4 = 0$
无名指	MCP	$0 \leq q_1, -0.2588 < q_2 < 0.7071,$ $-0.138 < q_3 < 0.138, -0.1951 \leq q_4 < 0.1951$
	PIP	$0 \leq q_1, 0 \leq q_2 < 0.866, q_3 = q_4 = 0$
	DIP	$0 \leq q_1, -0.0436 < q_2 < 0.7071, q_3 = q_4 = 0$
小指	MCP	$0 \leq q_1, -0.2588 < q_2 < 0.7071,$ $-0.1531 < q_3 < 0.1531, -0.2164 \leq q_4 < 0.2164$
	PIP	$0 \leq q_1, 0 \leq q_2 < 0.9239, q_3 = q_4 = 0$
	DIP	$0 \leq q_1, -0.0436 < q_2 < 0.7071, q_3 = q_4 = 0$

在表 3-2 中给出了手部关约束转化为关节姿态四元数的数值约束结果, 其中 $q_i$ 对应公式 3-14 中的 $h_i(\mathbf{q}_{L1}, \mathbf{q}_{L2})$ , 也即为系统状态的非线性约束, 在得到如表中所示的等式约束集合和不等式约束集合之后, 将其代入式 3-11 中, 之后利用 Active-set 方法进行约束求解, 即可得到符合生理约束的手部姿态。本文所用卡尔曼模型系统状态为各惯性传感器的姿态, 生理学上的关节角约束无法直接作用在系统状态值上求解, 因此将相邻骨节姿态差异视作手指两骨节所形成关节的姿态, 将手部的关节角约束转化为基于四元数的数值等式与不等式约束集合后, 得到关于基于骨节姿态值的等式与不等式函数集合, 利用投影法将针对系统状态的约束集合融入卡尔曼滤波系统中, 将关节姿态的约束问题

转化为系统状态的约束集合求解问题，在约束求解中，本文利用 **Active-set** 方法约束求解，最终得到符合约束的系统状态。

### 3.3.3 非线性约束下的姿态求解

本文将关约束融入到卡尔曼滤波系统的状态方程中，针对状态方程中约束问题求解，需要定义最小化约束的目标函数，目标函数用于衡量无约束状态下手部骨节姿态与符合约束的手部关节范围的接近程度，目标函数的设计可以使用欧氏距离作为目标函数的度量方式。对于每个手部骨节计算无约束状态下的姿态与约束范围之间的欧氏距离，将所有关节的距离进行加权求和，得到一个总体距离值作为目标函数的输出，用于度量无约束状态与约束面上的点之间的差异，目的是最小化这个差异，以找到与无约束状态最接近且符合手部关节范围约束的姿态，用  $\tilde{\mathbf{x}}$  表示。对于  $\tilde{\mathbf{x}}$  的求解运算，具体可以描述为如式 3-15 不等式约束优化问题。

$$\begin{cases} \tilde{\mathbf{x}} = \operatorname{argmin}_{\tilde{\mathbf{x}}} (\tilde{\mathbf{x}} - \mathbf{x})^T (\tilde{\mathbf{x}} - \mathbf{x}) \\ h(\tilde{\mathbf{x}}) = 0 \\ g(\tilde{\mathbf{x}}) < 0 \end{cases} \quad (3-15)$$

在式 3-15 中， $\tilde{\mathbf{x}}$  表示符合约束方程组的手部姿态最优值，也即满足约束关系的系统状态， $\mathbf{x}$  则表示无约束状态下的系统手部姿态，求得符合约束并与当前系统状态欧氏距离最短的手部姿态， $h(\tilde{\mathbf{x}})$  表示关节姿态四元数满足的等式约束集合， $g(\tilde{\mathbf{x}})$  表示关节姿态需要满足的不等式约束集合。

针对最小化目标函数的问题，存在多种最优化算法，包括梯度下降、共轭梯度法和牛顿法等。选择适当的算法取决于问题的特性以及约束条件的复杂性。对于简单的约束条件，如角度限制或线性约束，可以使用基于梯度的方法进行求解。然而，对于更为复杂的约束条件，例如非线性约束，可选择其他优化算法，如牛顿法和 **Active-set** 算法进行求解。在涉及非线性约束的最优化问题中，除了线性规划算法，还可采用其他数学优化方法。具体而言，牛顿法是一种典型的非线性规划算法，它通过利用二阶泰勒展开来逼近目标函数和约束函数，将原问题转化为一系列二次规划子问题。通过迭代求解这些子问题，牛顿法能够快速收敛到条件的最优解。另一方面，**Active-set** 算法按照约束的工作状态将其划分为工作集与非工作集，以确定迭代搜索的方向。综上所述，对于包含复杂非线性约束的问题，选择专门的非线性规划算法有助于获得更高的计算效率和更精确的最优解。

本文综合考虑采用 **Active-set** 方法进行约束求解，它是一种常用的数值优化方法，该方法通过迭代地将约束条件中的一部分视为等式约束，并通过求解相应的等式约束子问题来逼近原始非线性约束问题的最优解。根据上述不等式约束优化问题可知，在求解系统状态最优解时，需要将系统状态的约束按照等式约束和不等式约束进行分类预处理，具体分类步骤以及结果已经在上小节中给出，约束求解流程简图如图 3-6 所示。

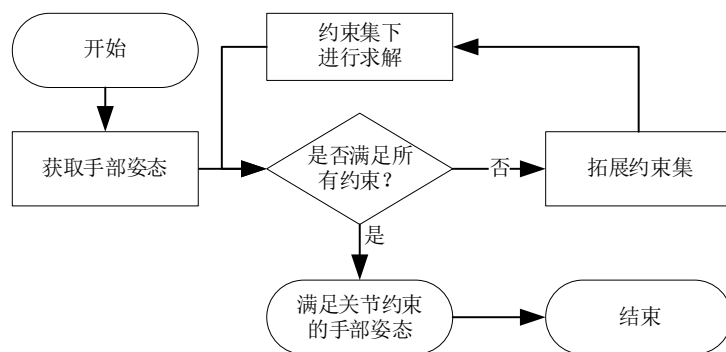


图 3-6 进行手部姿态约束求解的流程简图

在图 3-6 中, 将当前无约束的手部姿态 $\mathbf{x}$ 作为初始状态, 初始状态下, 判断当前无约束手部姿态是否符合所有的手部约束, 如果满足所有的约束信息, 则不需要进行约束求解, 若存在手部姿态不满足的约束条件, 则在每个迭代步骤中, 将系统状态不符合的约束条件添加到约束集合中, 并将其转化为等式约束, 就可以得到一个等式约束子问题, 其中目标函数与原始问题相同, 约束条件仅包含等式约束。通过求解等式约束子问题, 可以得到一组候选解, 表示在无约束情况下的手部关节姿态。在获得等式约束子问题的解之后, 需要检查候选解是否满足所有的关节姿态约束。如果候选解违反了某些约束条件, 将其添加到约束集中。重复执行上述步骤, 直到候选解满足所有约束条件且约束集不再变化或者达到最大迭代步数为止, 这表明已经找到了满足手部关节姿态约束条件的最优解 $\tilde{\mathbf{x}}$ 。

### 3.3.4 实验分析

本实验为手部姿态误差计算与对比实验, 目的在于探究优化方法的准确度, 与未优化方法相比是否降低了误差。鉴于使用角度尺以及组合测量装置作为测量手指关节角度的测量装置时需要研究人员记录角度值, 研究人员的操作会导致姿态的变化以及读数的不稳定, 而光学设备在环境光线充足条件下具有较高的准确度, 因此本实验使用商业手部姿态捕捉装置 Leap Motion 进行手部姿态捕捉, 将其作为手部的参考姿态。

#### 1) 参与者与实验环境

本次实验招募了 12 名参与者, 其中男性 9 名, 女性 3 名, 参与者年龄分布在 20 岁至 25 岁之间( $M = 22.08, SD = 2.11$ ), 惯用手均为右手。

实验所用计算机的显示器为 27 寸 2560\*1440 分辨率戴尔品牌, 交互设备为包含惯性传感器和信号处理单元的数据手套, 测量装置为 Leap Motion 运动捕捉装置。

#### 2) 实验任务与实验设计

在本实验中, 参与者需要根据提示做出不同的手势进行数据采集和对比。实验将伸展手势 (Open)、握拳手势 (Close) 和 Ok 手势 (Ok) 作为本次实验的三种手势, 手势示意图如图 3-7 所示。本次实验要求参与者右手佩戴数据手套依次执行伸展手势、握拳



手势和 Ok 手势,每种手势执行 10 轮,每次在抓握姿态稳定后采集 10 秒钟并保存数据。由于 Leap Motion 装置通过视觉的方式进行姿态捕捉,因此需要本实验环境保持光源稳定,执行动作时用户运动范围一致,掌心保持面向镜头,并且在执行手势时小臂与腕关节保持稳定。

选择这三种手势是由于伸展手势和握拳手势往往是人手关节日常运动的最大与最小范围,使用这两种手势能够体现出姿态估计方法在边缘范围下的准确性,而 Ok 手势则是日常生活中常用的具有语义的手势,其表现为食指与拇指自然弯曲,其余三指自然伸展,选取该手势能够体现出用户执行有语义手势时的准确程度。

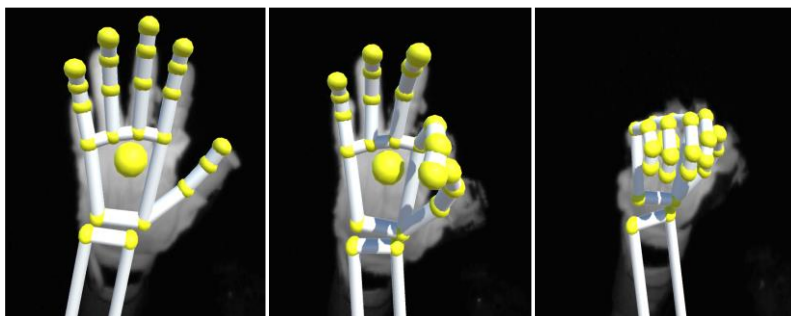


图 3-7 伸展手势、Ok 手势和握拳手势示意图

实验变量: (1) 姿态估计方法: 优化的手部跟踪方法、未优化的手部跟踪方法。  
(2) 手势: 伸展手势、握拳手势和 Ok 手势。

实验类型: 组内设计。

在通过实验得到用户执行任务过程中的数据后,对数据进行预处理操作,例如 Leap Motion 在进行手势捕捉时出现了误判左右手、数据采集过程中由于佩戴时的剧烈运动导致传感器桥接处断连等情况,本文根据实验时的标注及对结果的分析,采用插值法对误判左右手的数据进行替换。将 Leap Motion 的手部关节姿态值作为手部姿态参考值,利用传感器数据分别通过优化后的方法和未优化的方法进行姿态估计,将两种方法所得到的姿态值与 Leap Motion 所得姿态对比,计算姿态误差。

#### 4) 实验过程

实验开始前要求参与者坐在电脑显示器前,在实验人员的帮助下佩戴好数据手套并熟悉了解任务流程,帮助用户调整手臂合适的高度,参与者自然执行伸展、握拳和 Ok 手势,调整坐姿注视显示屏幕,此过程直到参与者能够表示熟练手势动作后结束。参与者的初始状态为手掌自然张开,当参与者听到执行手势时,执行相应的行为,系统开始采集参与者佩戴的数据手套数据以及 Leap Motion 的姿态数据,当数据采集 10 秒钟时研究人员告知用户本轮结束,每一轮间隔 30 秒,此过程一直重复直到实验完毕。

#### 5) 实验结果



通过实验获得了用户三种手势下的传感器数据和 Leap Motion 的姿态数据, 首先本研究将传感器数据通过未优化手部跟踪算法和优化后手部跟踪算法得到两种估计姿态, 之后将 Leap Motion 姿态和两种估计姿态的坐标系统一化, 并求两种方法的误差。

本文展示食指 PIP 在三种手势下优化后方法与未优化方法的误差对比结果, 如图 3-8 所示。在图 3-8 中由上而下分别为误差姿态四元数的值,  $q_1$  越接近 1 表示误差越小, 其余四元数值越接近 0 误差越小。

通过图 3-8 发现, 融合约束方法的对于食指 PIP 关节姿态求解优于未融合约束方法姿态, 为此本文针对各个参数进行了单因素方差检验(ANOVA), 判断两种方法的姿态误差是否具有显著差异。进行参数检验时样本需要满足正态分布, 针对两种方法下各个关节姿态误差进行正态分布检验, 本文使用对齐秩变换(ART)<sup>[66]</sup>对不符合正态分布的数据进行转换处理。结果显示食指 PIP 的参数  $q_3$  ( $F_{1,33} = 396.4$ ,  $p < 0.001$ ,  $\eta^2 = 0.46$ ) 和  $q_4$  ( $F_{1,33} = 396.3$ ,  $p < 0.001$ ,  $\eta^2 = 0.46$ ) 存在显著差异, 即两种方法对误差的影响显著。而姿态参数  $q_1$  的误差均值在两种方法之间并未表现出显著差异 ( $F_{1,33} = 1.724$ ,  $p = 0.1936$ ), 而姿态参数  $q_2$  的误差均值在两种方法之间并未表现出显著差异 ( $F_{1,33} = 0.065$ ,  $p = 0.7989$ )。该结果表明通过引入手部约束对手指弯曲方向的变化影响并没有表现出显著差异, 但降低了佩戴错位问题导致的其他旋转轴的误差。



图 3-8 食指 PIP 在三种手势下的姿态误差

通过计算发现, 优化方法对姿态系数  $q_3$ 、 $q_4$  的误差均值小于未优化方法的结果, 为了进一步分析误差离散程度与对比, 绘制姿态误差箱线图, 如图 3-9。

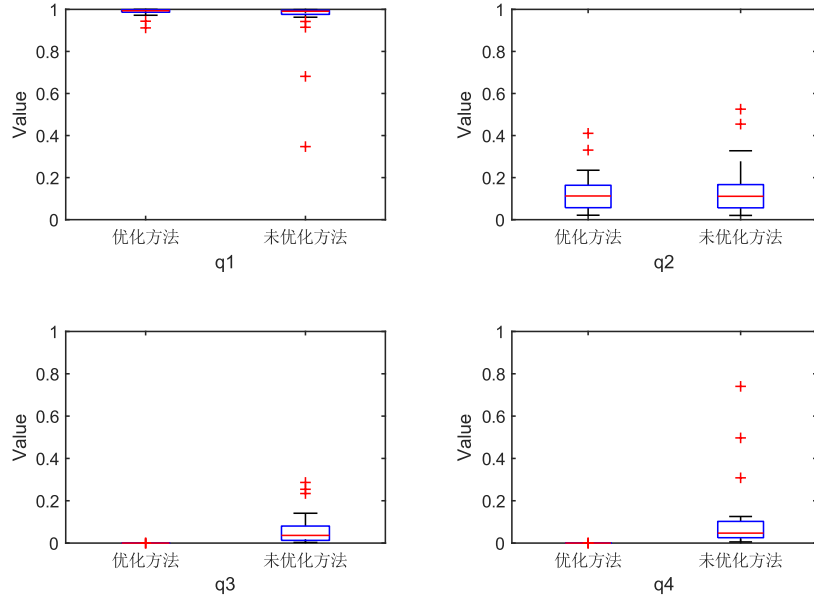


图 3-9 食指 PIP 姿态误差箱线图

在图 3-9 中，优化后方法的误差分布更加集中， $q_3$ 、 $q_4$ 相较于未优化方法的误差更接近 0 值，根据关节特性可知，PIP 只具有弯曲方向的自由度，即其姿态四元数参数 $q_3$ 、 $q_4$ 的值为零值，然而运动引起的错位误差，则会给估计所得姿态引入其他方向分量的误差，使得传感器与骨节之间发生偏转，导致姿态出现误差。在全手姿态跟踪中，引入生理学约束降低了佩戴错位误差导致的姿态估计误差，优化方法相较于未优化的全手行为跟踪方法，能够获得更真实的手部姿态。

本文设计的手部姿态测量实验，优化后方法在求解 PIP 姿态尤其在参数 $q_3$ 、 $q_4$ 的求解上表现出更好的准确性，误差显著低于未优化方法。手指 PIP 只有一个自由度，即只可以进行弯曲动作，不能进行旋转等其他动作，姿态四元数参数 $q_3$ 、 $q_4$ 值为 0，能够真实的反应出手部的姿态信息。综上所述，融合约束方法在手部姿态求解中优于未优化方法，能较好地降低由于运动错位导致的估计误差问题，在实际应用中，融合约束方法可以在手势交互领域作为一种更优的选择。

### 3.4 本章小结

本章研究了一种全手行为跟踪优化方法，通过佩戴数据手套进行手部跟踪，通过实验表明，结合约束的全手跟踪优化方法相较于未优化方法，在 PIP 姿态四元数两个分量的误差具有显著差异。与现有方法相比，不需要使用额外的模态信息，可以得到符合真实手部结构手部姿态结果，降低了可穿戴设备中佩戴错位的问题，这为接下来的基于姿态的手势交互技术提供了稳定的姿态信息。

## 第4章 基于姿态的手势交互技术

本章旨在探索并设计以用户为中心的自然手势交互方式,包括用户参与的交互手势集设计、构建映射模型和基于手部震颤特征的交互意图识别三个方面。现有研究的手势集无法满足本节基于姿态的手势交互需求,且研究者构建的手势集合可能存在偏见问题,因此本章首先基于以用户为中心的设计方法,向用户解释手部姿态的含义,采用大声思考的方法记录用户的想法,最终建立基于姿态的交互手势集合。

基于用户构建的交互手势集合,本章探索了一种坐标映射方法和基于姿态的手势识别方法。针对用户提出当使用手掌的偏航转动控制光标时转动角度范围小,无法进行长距离移动的问题,本节通过构建长短距映射模型实现光标的移动控制,发掘手势集合手指关节功能,构建决策树进行手势分类。

在手势交互领域中,交互系统功能与用户交互意图不匹配的问题一直存在,例如用户交互过程中活动手腕,交互系统会持续将用户行为转化为交互指令而出现的误识别问题,也被称为空中手势交互的 *Midas Touch* 问题。交互系统可以通过获取用户的交互意图提高交互的自然性和流畅性。在执行手势的过程中,用户行为由用户主观意图驱动并具有肌肉震颤信息。通过研究发现,健康人群的震颤信号是一种线性随机过程,而用户的交互意图则是一种用户自发的主观行为,具有非线性特性,两者通过神经元传达到肌肉发生作用,因此探究震颤信号中偏离线性的非线性信息,为识别用户的交互意图提供了一种解决思路。本章依据震颤信号的特征进行分析,设计用户实验采集用户行为的数据,提取震颤信号的特征以识别用户是否具有交互意图和探究交互意图强弱的影响。

综上所述,本章首先分析平面交互的基本功能,再通过用户实验设计基于姿态的手势集合。其次针对构建的手势集合研究了一种坐标映射方法和手势识别方法。最后分析健康人群的肌肉震颤与交互意图的联系,设计实验评估不同行为下的震颤特征,并通过多种分类器进行分类,结果表明利用肌肉震颤判别交互意图具有可行性。

### 4.1 基于姿态的平面交互手势设计

本节主要内容为建立以用户为中心的平面交互系统的手势集合。手势交互一般包括 2D 平面交互和 3D 空间交互,2D 平面交互是指用户与二维平面上的信息进行交互,典型的 2D 平面包括智能手机、大屏交互界面等,用户通过执行手部动作操作平面中的对象。3D 空间交互则是针对三维虚拟空间进行交互,通过捕捉手部姿态,构建虚拟手实现与虚拟空间中的物体交互,典型 3D 空间交互包括 VR 交互、3D 游戏等。本研究针对平面手势交互进行设计,第一部分在平面交互中,将平移、选择、抓取、旋转和缩放作为平面交互的基本功能。第二部分通过用户参与式设计方法,收集用户自然手势和偏好,设计出一套以用户为中心的基于姿态的手势集合。第三部分探讨了该手势集合的实现方

法，包括空间坐标映射与指令转换两部分，实现了一种符合用户偏好并支持多种典型操作功能的手势交互方法。

#### 4.1.1 平面交互基本功能

本小节探讨在平面手势交互中常见的功能与逻辑，将平移、选择、抓取、旋转和缩放作为本研究平面交互的基本功能。

在平面交互中，由于只有二维空间的限制，手势设计的自由度相对受限，因此选择恰当的基本功能尤为关键。本文从手势交互的相关研究中<sup>[67]</sup>，总结了平面交互的核心功能：移动控制、定位与确认对象、变形控制三个部分。本研究将平移、选择、抓取、旋转与缩放作为平面交互的核心功能，其能够有效覆盖平面交互的主要场景与需求。其中平移和选择满足定位与确认的基本需求，抓取、旋转和缩放构成了对象变形控制的主要形式。它们覆盖面广，容易记忆与识别，同时遵循用户认知习惯，是平面单手空中交互设计的最佳选择，基本功能的功能描述与场景见表 4-1。这五类基本功能可以涵盖用户在平面界面操作时的主要需求，也为基于手势的平面交互建立了控制基础。后续工作将在此框架内对特定的手势体系进行优化和评估，以丰富平面空间场景下基于手势的交互方式。

表 4-1 五类基本功能描述与场景

交互功能		功能描述	场景
平移	上	向上移动	在平面空间中，控制光标/对象进行上下左右的移动。
	下	向下移动	
	左	向左移动	
	右	向右移动	
点选		选择对象	选择平面空间中的对象，后续执行抓取、旋转、缩放手势。
抓取		抓起对象	抓取平面空间中的对象，移动对象位置到目标区域。
旋转	逆时针	使对象逆时针旋转	对平面空间中的对象进行旋转，旋转到目标状态。
	顺时针	使对象逆时针旋转	
缩放	放大	使对象放大	调整对象的大小到目标状态。
	缩小	使对象缩小	

在表 4-1 中，针对平移功能，用户需要频繁进行界面内的位置变换操作，如平移地图或图片，因此设计平移手势成为了手势交互的基础功能。选择功能可以让用户精确地选择特定目标，是大多数交互的基础。抓取功能模拟的手部抓握动作，可实现对象的抓取和移动。旋转功能通过手掌的旋转来控制对象的方向变换。最后，缩放功能允许通过手势控制界面或对象的显示比例。

尽管先前研究构建了各种基于手掌和手指姿态的手势集<sup>[67]</sup>，经过对比考虑，本研究

认为仅依靠手部姿态难以有效实现平面交互所需的这些功能,尤其平移功能无法得到很好的映射,例如用户执行平移向左的功能时,只需要将手从左边移动到右边,在这个移动的过程中,用户的手部姿态不会发生变化。

此外,完全依赖用户记忆预设姿态的交互方式,存在较高的认知和使用难度。为充分发挥手势交互的直观性与易用性优势,本研究将采用用户参与的方法,设计实验构建手势集。

#### 4.1.2 用户设计手势实验

本小节设计了一项用户参与式的设计手势集合实验,实验目的为设计一组以用户为中心且基于姿态的空中手势集合。面向用户自主设计手势的启发式实验,最大的优点在于能够挖掘人类手部运动的自然映射习惯,获得高度符合用户直觉与认知的手势体系。统计研究显示,此类手势的可理解性和可记忆性明显优于研究者自顶向下预设的手势集。此外,面向用户的手势设计还能够显著降低交互记忆与认知负荷,减少用户的训练时间,提升体验流畅度。本研究通过让普通用户根据操作需求自由设计手势,获得了更加可用与人性化的平面交互手势体系,为后续的交互系统改进和优化奠定了基础。

##### 1) 实验参与者

本次实验招募了7名参与者参与实验,其中5名男性和2名女性,年龄在23岁至25岁之间( $M=23, SD=1.5$ ),其中6人惯用手为右手。

##### 2) 实验任务和设计

本实验试验任务为用户观看功能演示动画并进行手势设计。本实验采用了启发式的研究方法,即通过功能动画的示意和解释,激发用户自主设计手势的思维,用户设计手势场景如图4-1所示。

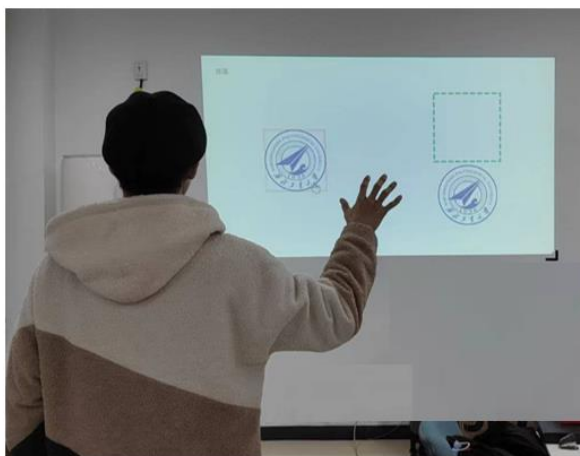


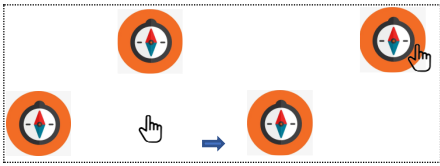
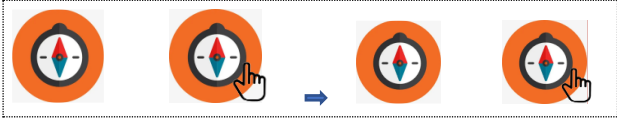
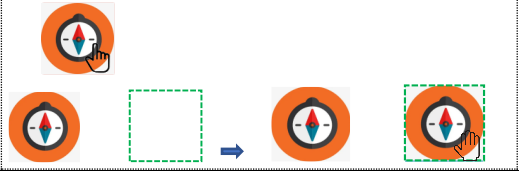
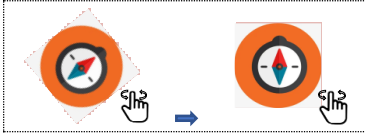
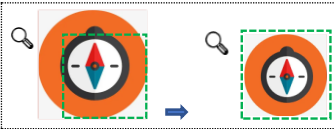
图4-1 用户设计手势环境图

实验中向用户展示平面空间中五类功能的动画,包括平移、选择、抓取、旋转和缩放,并解释其含义、应用场景和设计要点,作为手势设计的参考指引。要求用户采用大声思考法<sup>[68]</sup>,即在设计手势时将想法大声讲出自己的想法,研究人员通过录制和分析用

户的言语数据，挖掘手势设计的认知逻辑，建立用户思维模型。

为了让用户更加直观的感受功能，研究人员绘制了每个功能的动画，见表 4-2，并向参与者描述每个功能的作用和效果。研究人员告知用户需要完全基于姿态进行交互的前提下，令用户针对平面交互的五种功能进行手势设计，其中某些功能对应多个子功能，平移分为上移、下移、左移、右移，旋转分为顺时针旋转、逆时针旋转，缩放分为放大与缩小，实验全程记录用户的设计手势与思考内容。在表 4-2 中，示意图箭头左侧表示操作前的对象信息，箭头右侧表示操作完成后的结果。

表 4-2 平移、点选、抓取、旋转、缩放

功能	功能效果示意图
平移	
点选	
抓取	
旋转	
缩放	

为了探究手势集合的一致性，本研究中引入手势集合的一致性系数<sup>[69]</sup>对手势集合进行分析，手势的一致性系数定义如公式 4-1。

$$AR(r) = \frac{|P|}{|P|-1} \sum_{P_i \in P} \left( \frac{|P_{i1}|}{|P|} \right)^2 - \frac{1}{|P|-1} \quad (4-1)$$

在公式 4-1 中， $P$ 是参与者的数量， $P_i$ 是选择手势 $i$ 的人数。 $AR(r)>0.5$  表示具有非常高的一致性， $0.5>AR(r)>0.3$  具有较高的一致性， $0.3>AR(r)>0.1$  具有一致性。一致率越高，代表更多的参与者提出了相同或高度相似的手势提议，参与者对于执行某个功能的手势理解存在更高的一致性，该手势具有更好的可猜测性，从用户体验角度来看更自然、更容易被接受，获得了更高程度的用户共识。



### 3) 实验过程

实验开始前,向用户说明实验目的,要求用户填写知情同意书与基本信息调查问卷。之后向用户解释基于姿态的手势交互方法,明确告知用户通过手部姿态进行功能控制与手势设计,手部姿态即手部姿态包括手掌的俯仰动作、翻滚动作、收展动作,所有手指的关节角度以及所有关节的旋转过程,不能通过手的空间位置进行手势设计。研究人员向用户解释需要设计的五类功能,要求用户在大屏幕前使用单手进行空中手势设计,并告知用户在设计过程中,将自己的思考过程大声说出来,对任务的理解、手势选择的顾虑、怎样思考的、为什么这么设计等。每当用户设计完手势后,需要对自己设计的手势填写两个 7 点 Likert 问题:你是否认为该手势容易执行?(1:非常困难,7:非常容易);你是否认为该手势能够与该交互功能直观的联系到一起?(1:非常困难,7:非常容易)。当用户提出一组完整的手势集合之后,对用户进行访谈,探究的问题主要在于用户提出的手势集合中,给出最容易、最难和最满意的手势,并要求用户给出原因。最后再重新复述一边实验目的,实验完毕。

### 4) 实验结果

实验总共收集了 90 个手势,根据收集到的手势分析特征,对具有相同特点的手势合并为同一类手势,例如选择功能中,有用户提出五指张开的情况下食指弯曲表示点击,和用户提出其余手指并拢的情况下食指弯曲表示点击,这两种动作的核心都是食指的来完成点击功能,即将用户想要表达的手势核心动作如果一致,则可以归为同一类手势。

通过收集到的手势集合,根据公式 4-1,分别计算各个功能的一致性分数,手势集的平均一致率为 0.35,体现出手势集合具有较高的一致性,结果如图 4-2 所示。

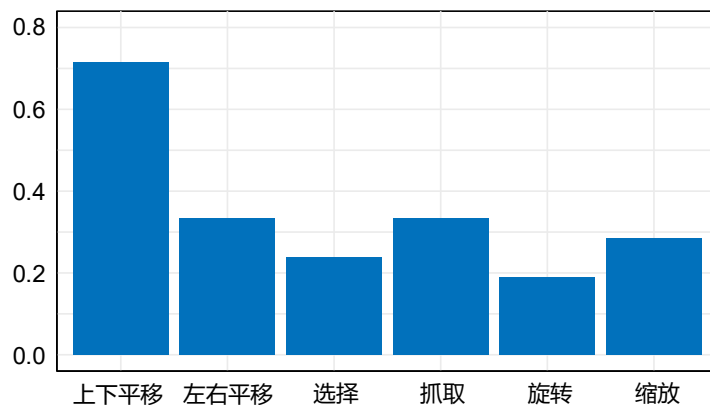


图 4-2 用户参与设计手势的一致率分数

在图 4-2 中,用户在光标上下平移的功能上,具有较高的一致率  $AR(r)=0.71$ ,只有一名用户使用手掌俯仰和四指弯曲表示移动,其余用户均选择手掌俯仰控制光标上下移动。左右平移和抓取功能的一致率分数均超过了 0.3,这两类功能具有较高的一致性,选择、旋转和缩放功能的一致率分数也超过了 0.1,这两个功能的手势集合具有一致性。

用户在进行光标上下平移的手势设计时,均提出了使用手掌俯仰的转动进行控制,

有一位用户表示“我想根据指尖的方向进行判断”，将指尖位置和光标位置联系起来。唯一不同手势的用户则是提出了触发式的方案，将手俯仰旋转和四指弯曲组合在一起，当检测到该动作时，光标以匀速移动，直到用户将手伸直放平。用户在光标左右平移功能上，出现了不同的结果，有四名用户选择了手掌偏航转动动作控制光标左右平移，有一名用户说“我的手偏航转动的时候，活动范围有点小，不像手俯仰转动那样大范围”，于是放弃了使用手掌偏航作为左右平移的光标控制，选择了使用手掌翻滚转动作为光标左右控制。还有一名用户则是使用拇指指向手心作为区分，使用手掌的俯仰进行控制。也有用户沿用了自己的设计风格，采用了触发式的方案，使用手掌翻滚和四指弯曲控制左右移动。

在选择这一手势上，用户几乎都没有过多思考就给出了手势，有三名用户使用食指点击进行选择，有一名用户说“这个手势就像点鼠标一样，点一下就选择了”，有两名用户使用食指和拇指捏住的手势，有一名用户说“选择可能会很频繁的触发，因此我选择食指与拇指捏住这个手势，因为这个很容易执行”，有两名用户选择握拳作为选择手势。

在抓取功能上，有五名用户选择了五指并拢，有一名用户一开始也选择了五指并拢，但是发现执行该手势时再执行移动手势受限，于是换成食指伸出，食指的方向控制抓起后的物体移动。同样有一名用户沿用了自己的设计风格，两指捏合的方式控制抓取，其余手指弯曲控制触发式的移动指令。

旋转手势中，有三名用户通过五指张开触发旋转指令，模拟抓住这个物体进行旋转控制，有两名用户则是将自己的手看作旋转的物体，执行翻滚动作控制对象的旋转，有两名用户采用触发式的指令进行控制，伸出食指和中指，分别通过偏航和俯仰进行旋转。

缩放手势中，有四名用户通过调整手指指尖的距离进行放大与缩小，有一名用户则是使用多个手指的，有一名用户采用手指和旋转来进行操作，拇指表示触发缩放，使用手掌翻滚来进行放大缩小的控制。

在用户提出手势之后，让用户立刻回答两个 7 点 Likert 问题，统计这五种功能的得分，绘制出箱线图如图 4-3 所示。

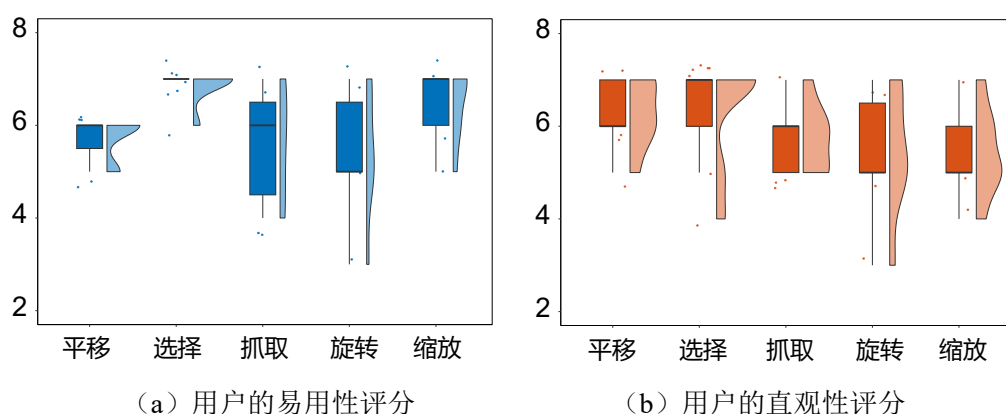


图 4-3 用户手势评分箱线图



在图 4-3 中, 在手势易用性评价中, 选择手势有六名用户打了 7 分, 一名用户打了 6 分, 都认为自己提出的选择手势容易执行。针对平移功能的手势则有用户提出, 在进行左右移动时手的旋转角度受限, 不容易控制光标移动到屏幕的边界。抓取手势在易用性上得分较低, 分析用户的想法得出, 用户认为执行抓取手势之后, 还要控制抓取的物体移动, 所以认为该功能较难执行。在手势直观性评价中, 用户 7 对自己提出的旋转手势打了 3 分, 该用户提出的手势为伸出食指和中指, 手掌俯仰控制旋转角度。

针对于不同的用户而言, 各自喜欢的风格可能也会有很大区别。有的用户喜欢模拟现实的操作, 用户 5 提出像点击鼠标一样使用食指点击执行选择功能, 将大拇指和食指指尖的距离看作物体的对角线执行放大和缩小的功能。有的用户并不喜欢将功能细分到具体的手指, 用户 3 说“我不喜欢细分手指的功能, 我觉得手势交互最好的就是我大范围的操作, 手抓一下能实现大部分功能就好”。有的用户不擅长使用连续的控制, 喜欢触发式的控制。用户 7 所提出的所有手势, 都是用手做出一个指令, 系统识别后启动对应的功能。用户还考虑到了执行频率的问题, 对于执行频率高的功能, 选择更简单的手势。

表 4-3 最终选择的手势集



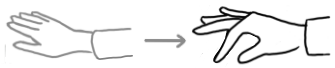
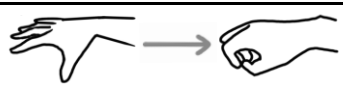



交互功能	动作描述	手势
平移	控制光标进行上/下移动	以拇指掌指关节与小指掌指关节为轴进行转动。 
	控制光标进行左/右移动	以手掌所在平面法向量为轴进行转动。 
点选	食指做出点击动作、余下手指自然伸展。	
抓取	五指自然握拳。	
旋转	控制目标逆时针/顺时针转动	五指弯曲, 以手掌所在平面法向量为轴进行转动。 
缩放	对目标进行放大操作	食指与拇指捏合、余下三指自然握拳, 食指与拇指自然张开呈 L 形。 
	对目标进行缩小操作	食指与拇指张开呈 L 形、余下三指自然握拳, 食指与拇指自然捏合。 

表 4-3 给出了本研究最终确定的手势集合与对应动作描述。本研究中用户自定义手势集的开发遵循最大群体优先的原则。具体而言, 对于每个操作功能, 通过统计分

析用户提交手势提案的相似度分布，选择相似度最高的手势组作为该功能的映射。这可以尽可能避免一个手势对应不同操作导致的命令冲突。当不同操作之间手势指令存在冲突时，赋予用户群体更大的子集自定义的手势以优先权，即多数人优先，这种方法能够在平衡操作一致性与用户倾向性之间取得最优解。

## 4.2 平面空间下手势交互实现方法

本节主要内容为探究以用户为中心，基于姿态的手势集合实现方法，主要分为坐标映射与手势指令实现两部分。全手行为跟踪方法是平面空间手势交互的核心部分，由于人手与交互空间维度的不一致性，在平面空间交互中还需要交互映射技术，将人体交互空间的手势行为转换为大屏空间坐标系的指令，其中关键是对手势空间和大屏空间的转换，本节分为两部分分别介绍系统所用的空间映射方法及手势集分类方法。

### 4.2.1 手部姿态与平面空间坐标映射

本小节主要介绍了基于手部姿态的平面交互映射模型。映射模型的目标是将用户手势角度变化映射为平面光标的位置控制。具体而言，以用户手部向量为基准坐标，测量用户单位时间内手掌俯仰和偏航运动的角度变化。其次将角度变化量映射为平面空间中光标竖直和水平方向的位移。这种映射方式参考了用户习惯，即使用手掌转动控制光标上下移动。为使控制更加自然连续，将中指指骨末端相对手腕的变化距离映射为光标变化距离。

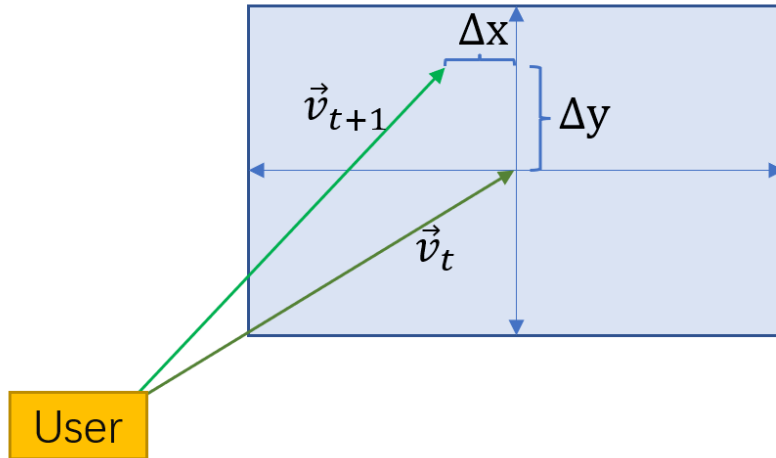


图 4-4 坐标系转换模型示意图

在图 4-4 中， $\vec{v}_t$  表示用户  $t$  时刻下的初始向量，而  $\vec{v}_{t+1}$  则表示利用手部姿态变化量旋转后的向量，当用户在单位时间内执行旋转动作，该过程用四元数  ${}_{H_t}^L \mathbf{q}$ ，执行偏航动作作用四元数  ${}_{L_t}^{H_{t+1}} \mathbf{q}$  表示，该转动过程的姿态四元数满足公式 4-2，用户手所在向量  $\vec{v}_t$  的转动结果如公式 4-3。

$${}_{H_t}^{H_{t+1}} \mathbf{q} = {}_{L_t}^{H_{t+1}} \mathbf{q} \otimes {}_{H_t}^L \mathbf{q} \quad (4-2)$$

$$\mathbf{v}_{t+1} = R(\mathbf{H}_{H_t}^{t+1} \mathbf{q}) \mathbf{v}_t \quad (4-3)$$

将公式 4-3 向量的  $\mathbf{v}_{t+1}$  的 x 分量和 z 分量的值映射到平面空间中竖直方向的移动和水平方向的移动。这种映射方式受到 4.1 节中用户的启发，用户选择手掌俯仰转动作为控制光标上下移动的手势时，提出根据指尖的位置进行控制。鉴于不同的手势需要手指运动实现，因此本研究将中指掌骨视为移动的向量，掌骨末端相较于手腕的空间移动，映射为光标的空间移动。

上述过程将手部姿态的变化角度映射为了平面空间中的移动距离。原始数据到设备控制指令的映射是实现手势交互的关键步骤。常见的基于原始数据的映射方法包括线性映射、非线性映射、基于模型的映射等。线性映射模型简单直接，但会引入失真。考虑到人手腕关节转动范围和运动时间特点，设计合理的阈值，实现光标位移的非线性平滑控制。这种模型使手势控制更加符合人机交互的自然特性，提升了交互体验。线性映射通过线性函数建立输入和输出之间的对应关系，该方法简单直接，但容易引入极端值或失真。非线性映射如指数、对数等复杂映射函数，可以提供更平滑自然的控制效果，但需要精心设计。基于模型的映射是指运用机器学习模型来学习输入和输出之间的复杂映射关系，可以自动适应不同用户的偏好，实现个性化交互，但需要收集大量训练数据。

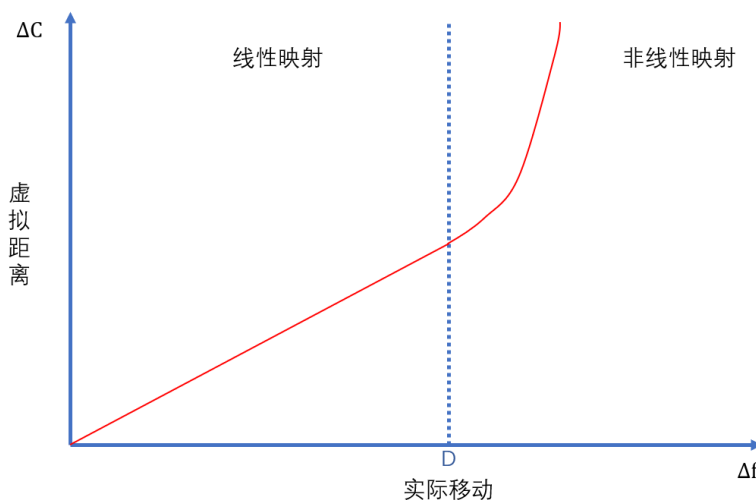


图 4-5 基于分段函数的非线性映射图

基于 4.1 小结中用户提出的手掌运动范围受限的担忧，本文改进了这一映射方案。有研究者<sup>[70]</sup>研究了一种名为 Go-Go 的交互技术，使用交互式增长方法满足交互的非线性映射。映射函数的设计在平衡映射的准确性、连续性、容错性等方面需要综合考虑。同时，评估和迭代优化也是确保映射质量的重要环节。

在图 4-5 中，横轴表示手部姿态旋转的角度，纵轴表示光标在平面中的位移距离。这表明映射函数为一段分段函数，研究表明<sup>[71]</sup>人手腕关节的俯仰方向活动范围为 110 度，偏航方向活动范围为 50 度，这为本研究设计分段函数的临界点与函数提供了参考。综合考虑，本研究建立映射分段函数，如公式 4-4 所示。

$$\Delta c = \begin{cases} a \times \Delta f, \Delta f < D \\ a \times \Delta f + (\Delta f - D)^2, other \end{cases} \quad (4-4)$$

在公式 4-4 中,  $\Delta c$  表示光标移动位置变化量,  $a$  表示映射系数, 对应于手掌转动幅度与光标移动位置映射系数,  $\Delta f$  表示掌骨相对于掌心移动距离变化量, 可通过公式 4-3 得出,  $D$  表示在单位时间内手指移动的距离临界值。根据该公式进行坐标映射, 实现了基于手部姿态的平面交互映射模型。

#### 4.2.2 基于手部姿态的交互指令实现

本小结主要介绍了 4.1 小节中用户提出手势集的识别方法, 分析在每种手势下手指的功能和状态, 建立决策树模型进行手势分类。分析 4.1 小结用户提出的手势集合, 发现可以将其转化为指令型的静态手势和动态的控制, 针对基于佩戴设备采集的手部和手指关节数据, 通过决策树将这类信息作为判断属性, 递归地构建出一个树形的手势分类器。

在用户提出的手势集合中, 平移手势仅需要手掌的偏航与俯仰操作, 不需要手指参与操作。点选只需要食指弯曲, 其余手指不需要弯曲。抓取动作则是执行握拳手势, 所有手指都弯曲到最大角度。旋转的手势更像模拟抓住了物体进行旋转, 通过手掌偏航控制旋转方向。缩放则是中指、无名指和小指弯曲到最大角度, 通过食指和拇指指尖的距离来控制缩放程度, 拇指和食指之间的距离可以利用关节角度通过前向动力学进行估算。为了更直观的表达五种手势中各个手指的参与情况, 本节总结了不同功能下手指参与的关节数目, 见表 4-4。

表 4-4 手指在不同手势中弯曲关节数

功能\手指	拇指	食指	中指	无名指	小指
平移	0	0	0	0	0
点选	0	1	0	0	0
抓取	2	3	3	3	3
旋转	1	2	2	2	2
缩放	0	0	3	3	3

通过表 4-4 可以发现, 仅通过判断食指的弯曲关节数, 就可以区分点选、抓取、旋转手势, 但在复杂的环境中误识别现象更容易发生, 因此综合多个手指考虑以提高手势识别模型的准确率。

本节通过构造决策树进行手势分类。决策树由于其模拟人类决策过程的特性, 被广泛应用于基于佩戴设备的静态手势识别任务中。在该任务中, 使用包含多种动作传感器的智能数据手套采集手部和手指的运动数据, 并人工标注对应的手势类别。其次利用决策树算法以手部运动特征为判断依据, 学习拆分规则, 建立决策树模型。该模型能够根据输入的手部运动数据, 从树根遍历到叶子, 得到手势的分类结果。决策树模型表现出

较强的解释性,通过树的结构和决策规则,能直观地理解计算机为何将一组运动数据识别为某个类别。同时,决策树也支持增量学习,当数据集增加时只需更新部分树分支,这些特性使其在基于佩戴设备的静态手势识别等领域中被广泛采用。根据手势集合特征构造的决策树结构图如图4-6所示。

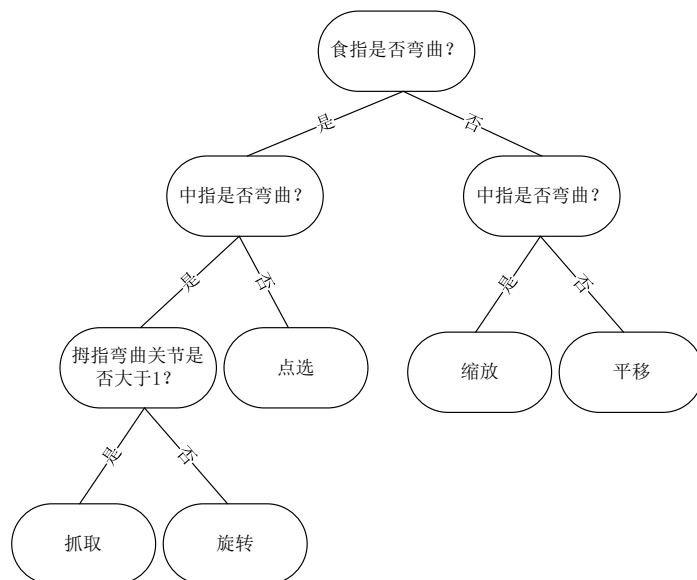


图 4-6 手势识别的决策树结构

在图4-6中,通过判断食指是否弯曲,则将手势集合划分为平移、缩放一类,点选、抓取、旋转一类。再通过判断中指是否弯曲,可以区分出平移和缩放、点选手势,由于抓取和旋转都用到了五根手指,因此需要判断关节数目来进行区分,在本决策树中选择了拇指进行判断,因为旋转手势对拇指只有一个关节弯曲,更容易进行分辨。

### 4.3 基于姿态的交互意图识别方法

本小结主要探究通过震颤信号识别用户行为是否具有交互意图,为空中手势交互的Midas Touch 问题研究一种解决方法。Midas Touch 问题是人机交互领域常见的问题之一,指当用户进行自然动作,改变手掌或手腕的位置和角度时,这些无明确交互目的而出现的姿态变化,也可能被交互设备所采集的传感数据误判为用户有意进行的操控手势指令。例如用户手掌的无目标晃动,就可能会被错误识别为某种具体的指向或选择动作,并导致执行未预期的操作。

本节主要探究人体震颤信号与交互意图之间的关系,从神经学角度分析了静息态震颤、动作态震颤及含交互意图的动作态震颤的特征。健康人群的震颤是一种线性随机过程,而交互意图则并不符合简单线性随机模型,具有一定的非线性性质,因此探究震颤信号的线性随机过程特征和偏离程度,可以分析用户行为是否蕴含交互意图。本节基于震颤振幅、频率等时间/频域特征等线性随机过程特点,构建了14个特征。在不同难度交互任务中采集参与者的各类震颤信号,利用这些特征可以区分含交互意图的动作性震

颤和不含交互意图的震颤信息。最后本节探究交互意图强度强弱与区分含交互意图的动作性震颤和不含交互意图的震颤的联系。本节旨在通过健康人群的震颤信号特征，分析用户的交互意图，以实现更自然的手势交互。

#### 4.3.1 手部行为震颤与交互意图

本小节主要描述了行为震颤的特征和交互意图对行为的影响，通过震颤信号来区分用户行为是否具有交互意图及强弱意图对识别结果的影响，以通过这种方式更好的理解用户交互意图。

神经学的相关研究将人体的震颤分为静息性震颤、动作性震颤，在实验分析中，令参与者执行具有交互意图的运动，提取过程中含有交互意图的动作性震颤（IAT, Intent Action Tremor），而令参与者不具备交互意图情况下采集静息性震颤（RT, Rest Tremor）和维持姿态的动作性震颤（PAT, Postural Action Tremor）。通过震颤的特性挖掘震颤信号的特征，通过震颤信号识别运动中是否蕴含交互意图。

震颤是指身体部位近似有节律的不自主运动，主要见于面部、下颌、上颚、眼睛、躯干和四肢。四肢震颤的典型表现为不规则、持续的运动，其频率和振幅可能发生波动。正常或增强的生理性震颤常呈现杂乱无序的宽频谱特征，而病理性震颤则更为规则、缓慢，噪声小，频谱中含有尖锐峰值。频谱分析可提供震颤信号在频域的组成信息：宽频谱表示多种频率成分贡献显著，集中峰值表示某单一主频率成分占主导地位。频谱特征可用于分析和判断震颤的类型、程度及状态。

现阶段，研究中主要有两种主要的震颤分类<sup>[72]</sup>，第一种是基于观察到震颤时身体部位的活动状态，第二种是基于潜在疾病或条件的病因。本研究面向健康人群，采用另一种分类方式，对震颤的分类为当相关肌肉没有被激活，身体部位得到充分支撑以抵抗重力时发生休息性震颤；当相关肌肉被激活时发生的动作性震颤，包括姿势震颤、动能震颤、等长震颤、意图震颤和任务特定震颤。本章基于该分类方式探究震颤特征与交互意图之前的联系。

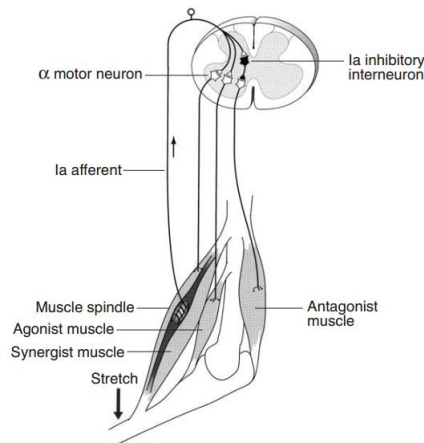


图 4-7 神经控制震颤信号发生图<sup>[72]</sup>

在震颤信号中,研究者 Elble 等人<sup>[73]</sup>指出正常震颤包含两个不同的振荡成分。第一个成分是肢体拉伸反射所形成的欠阻尼振荡,频率由肢体的质量和刚度决定,这一成分的产生源自肢体本身的被动力学特性。当外力或内力干扰肢体时引起的振荡,外加质量通常会降低频率,外加刚性会提高频率,该成分与肌电爆发无关,但当震颤加剧时,运动神经元的放电会通过反射性调节。第二个成分与肌电爆发直接相关。Deuschl 等人<sup>[74]</sup>表明静息性震颤频率通常在 4 Hz-6 Hz 之间,而姿势性震颤频率通常在 4 Hz-12 Hz 之间。研究者 Elble 等人<sup>[75]</sup>指出正常的生理震颤通常集中在 8 Hz-12 Hz 的范围,但不同人也可能存在较大差异,震颤是一种线性随机过程,具有时间反转对称性等特点。

而交互意图的形成是一个复杂的认知和决策过程,难以简单地用线性模型描述。具体来说交互意图的产生涉及复杂的大脑认知活动,会受到情感、经验、环境、目标等多种内外部因素的影响,这是一个多维的非线性过程。在交互前、中、后,用户的意图也会不断发生变化。这同样是个动态调整的过程,难以用静态的线性关系解释。不同用户由于其个体差异,在相似环境和目标下,会产生不同的交互策略,这种多样性也决定了非线性的特点。用户意图表达和意图理解中的语言、表情、手势等都是复杂的交互要素,互相影响也呈现出强非线性特点。

总的来说,交互意图的形成和演化是一个多源、动态、个体化复杂的过程,很难用简单的线性模型进行表达,应该视为一个非线性的现象加以研究和处理。这为通过震颤信号的线性特征差异分析交互意图提供了理论基础。

#### 4.3.2 交互中手部震颤提取方法

本小节使用巴特沃斯滤波器进行信号过滤<sup>[76]</sup>,保留 4 Hz-12 Hz 的震颤信号,根据震颤信号的振幅、频率和符合线性随机过程的特性,描述了震颤信号具有偏度和峰度特征和具有时间反演对称性,将这些特征作为特征集的备选以进行实验设计与分析。

##### 1) 震颤信号提取

根据时间序列可以对震颤进行测量,最简单和直接的是测量其平均振幅。如果信号平均值为零且整体无漂移趋势,那么振幅即为信号的均方根即标准差。为去除某些频率分量,可以对信号进行滤波处理,这相当于部分或全部去除其傅立叶变换,然后进行反变换生成新的滤波信号,如果存在整体漂移趋势,高通滤波可用于消除低频成分,从而更准确测量振幅。判断去除多少低频部分需要谨慎,例如呼吸、心跳频率影响位移记录,而震颤频率一般认为至少有 4 Hz,并且一般不高于 12 Hz。除振荡振幅外,从参考点偏离的总量也可反映病理学情况。位移信号可通过傅立叶分析分离漂移和震颤成分。滤波前后示意图如图 4-8 所示。



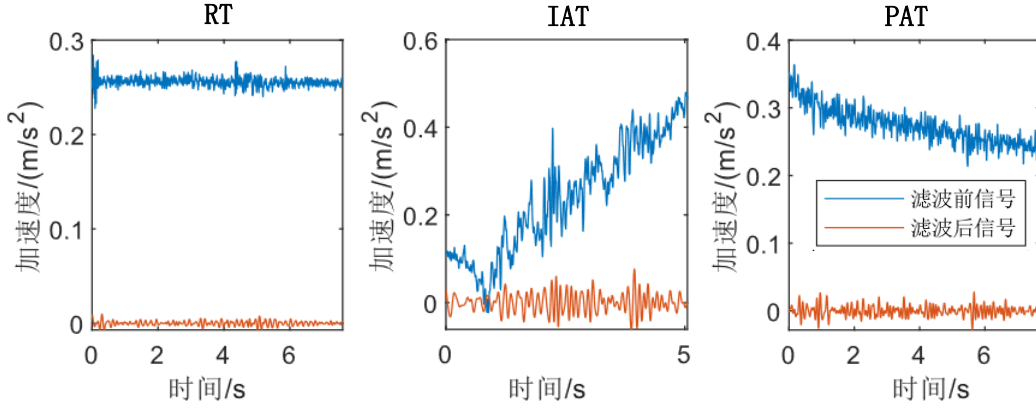


图 4-8 RT、IAT、PAT 的滤波前后信号

在图 4-8 中, 使用 10 阶巴特沃斯高通滤波器( $f_1 > 4 \text{ Hz}$ )进行带通滤波, 用于消除手部动作的低频信号, 再用 10 阶巴特沃斯低通滤波器( $f_2 < 12 \text{ Hz}$ )将高频噪声过滤掉, 保留静息性震颤信号(4 Hz-6 Hz)与动作性震颤信号(6 Hz-12 Hz)。通过这种方式能够去除运动的影响。综上, 时间序列分析可定量分析震颤的振幅、频率等参数, 振幅测量需要考虑信号预处理, 以提高测量准确性。

## 2) 线性随机过程特征

有研究者 Timmer J 等人<sup>[77]</sup>指出健康人的震颤信号是一种线性随机过程。而人的交互意图也是一种非线性的行为, 通过分析震颤信号的线性随机过程所具有的特征, 可以区分含有交互意图与不包含交互意图的行为震颤。

生理性震颤具有高斯随机性且基本过程为线性, 检测震颤是否含交互意图的简单方法是观察数据值分布是否偏离正态分布。常用方法是计算高阶矩, 这些高阶矩可以检测出信号分布与正态分布的偏差。例如第三阶矩(偏度,  $m_3$ )反映分布的对称性, 第四阶矩(峰度,  $m_4$ )反映分布尖峰情况。正弦波等规则对称振荡其偏度近似为零, 但峰度明显低于正态分布。

综上, 如果生理震颤为高斯随机过程, 可以通过计算高阶矩, 判断信号分布是否偏离正态分布, 以检测信号中是否存在额外的意向性成分, 这是一种简便有效的方法, 公式 4-5 给出了偏度的计算过程。

$$m_3 = \frac{1}{n} \sum_{i=1}^n \left( \frac{X_i - \bar{X}}{s} \right)^3 \quad (4-5)$$

公式 4-5 为偏度计算公式, 其中  $m_3$  表示偏度信息,  $n$  表示数据总数,  $X_i$  是样本中第  $i$  个观察值,  $s$  表示样本的标准差,  $\bar{X}$  表示样本均值。偏度是描述概率分布形状对称性的一个统计量。它反映分布曲线相对于中心的左右尾部延伸的情况。偏度为正表示右侧尾部更长, 负偏度则表示左侧尾部更长。偏度绝对值越大表示分布曲线越不对称。零偏度对应的是对称分布, 如正态分布。判断分布的偏斜方向和程度, 偏度是最有效的度量标准, 可判断实际分布与理论对称分布的差异。



峰度描述概率分布曲线的尖峭程度,反映尾部与峰值区域曲线形状特征。正态分布峰度为3,称为正常峰度。当实际分布的峰值区域相对较尖峭时,峰度大于3称为尖峭峰度。当峰值区域较为平缓时,峰度小于3称为平坦峰度。

$$m4 = \frac{1}{n} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{s} \right)^4 - 3 \quad (4-6)$$

公式4-6为峰度计算公式,其中 $m4$ 表示峰度信息, $n$ 表示数据总数, $x_i$ 是样本中第 $i$ 个观察值, $s$ 表示样本的标准差, $\bar{x}$ 表示样本均值。正态分布曲线代表了大多数实际分布均匀和对称地延伸到两个方向的尾部区域,与其峰值区域契合,这是一种典型和标准的分布形状。当实际分布曲线尾部相对于峰部突出时呈现较尖峭的非对称形状,如指数分布,其峰度大于3。反之,当实际分布曲线整体较为平缓时呈现钟形,其峰度小于3。通过与标准正态分布的比较,峰度直观反映分布整体形状的差异特征,是判断分布形状的重要指标之一,针对震颤信号的正态分布拟合如图4-9所示。

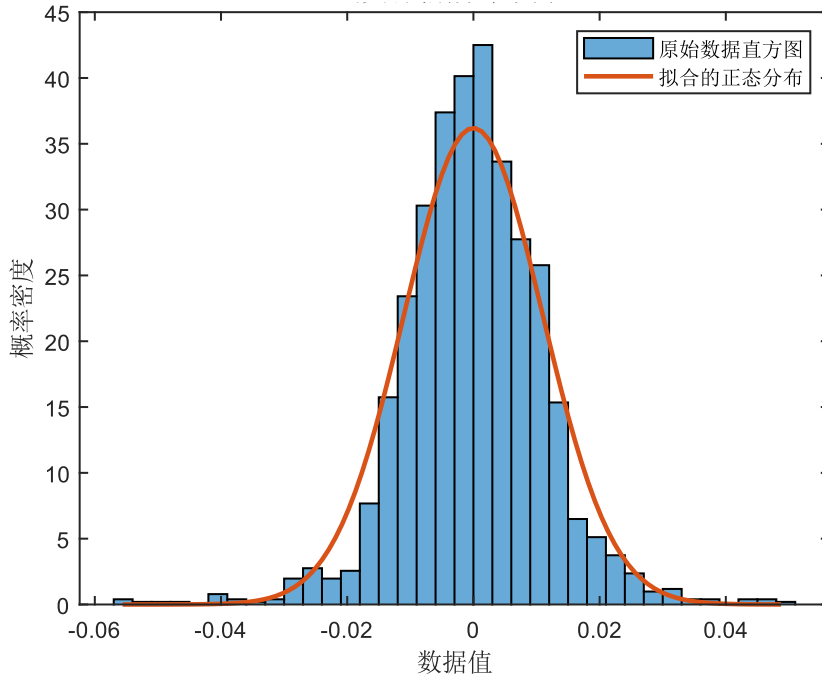


图4-9 健康人群震颤信号分布直方图

研究者 Edwards 等人指出<sup>[78]</sup>,震颤信号具有时间反转对称性,即在时间反转下性质不变,而非线性过程的性质通常会改变,也可以通过这种方法来判断是否蕴含交互意图的非线性差异。时间不对称性的一般测量可以用来诊断正确的非线性类型。其中一种方法是从相似值分别预测向前和向后的值(具有相同滞后),并比较两者差异的大小。对所有可能的起始值均进行这样的预测,将方差相加即可得到一个滞后相关的时间不对称性度量。时间不对称性的另一种度量方法是计算点和滞后点之间差异的立方和,见公式4-7。

$$R = \max \left\{ \frac{\sum_i (x_i - x_{i-j})^3}{(\sum_i x_i^2)^{3/2}} \right\} \quad (4-7)$$

在公式 4-7 中,  $j$  为时间滞后参数,  $X_i$  为样本中的第  $i$  个值。线性高斯过程具有时间反转对称性, 而非线性过程不具有该特点, 通过计算时间序列在正反两个时间方向上的预测误差, 可以判断时间不对称性, 从而识别信号的非线性成分, 实现对含有交互意图行为的非线性过程的有效检测。

如上所述, 震颤信号的线性随机过程所具有的特征可以有效判断信号的线性与非线性特性。频谱分析的常用方法还包括功率谱密度分析和主频率分析。功率谱密度反映信号功率在频域上的分布情况。通过离散傅里叶变换可以计算得到静息态和运动态震颤信号的功率谱密度值, 作为判断震颤强度的指标。主频率分析即识别信号中占主导地位的频率成分。通常震颤信号会包含一个主要的频率成分, 这可以反映震颤的类型和状态。短时傅里叶变换的时频分析可以有效提取该主频率。因此, 频谱分析中功率谱密度和主频率等参数的计算, 可以从频域角度更全面地分析判断震颤的特征。

#### 4.3.3 基于震颤的交互意图识别

本小节描述通过震颤特征进行交互意图的识别, 主要分为识别震颤信号是否用户行为蕴含交互意图以及交互意图强弱。首先介绍用户震颤信号的分类, 有无交互意图的震颤信号类型, 用于通过震颤区分行为是否具有交互意图。然后描述构建震颤的特征集合、对特征进行标准化处理、相关性检验去除高相关特征、使用重复 5 折交叉验证进行结果评估、使用准确率、召回率、F1 分数作为指标进行结果分析。最后设计用户实验进行分析。

根据 4.3.2 小结内容, 在本章中将用户有意图的行为所产生的震颤称为意图性动作震颤 (IAT, Intent Action Tremor), 将用户没有交互意图的维持动作产生的震颤称为维持性动作震颤 (PAT, Postural Action Tremor), 而用户肌肉没有被激活时无交互意图的震颤称为静息性震颤 (RT, Rest Tremor), 三种震颤信号原始数据见图 4-10。

在图 4-10 中, 最左侧为用户肌肉没有被激活时的震颤信号 RT, 震颤幅度较低。中间为用户执行交互任务, 具有交互意图的运动性震颤 IAT。右侧为用户维持姿态的运动性震颤 PAT, 这个过程用户不具有交互意图。

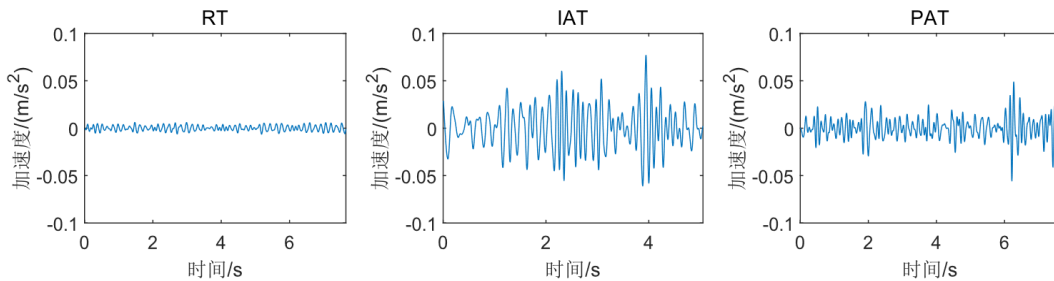


图 4-10 RT、IAT 和 PAT 震颤信号波形

这三种震颤中只有 IAT 具有交互意图,于是根据交互意图的强烈程度,设计不同的交互任务将对 IAT 进行进一步区分,判断用户所产生的 IAT 信号强烈程度,在不同交互意图情况下,用户震颤信号如图 4-11 所示。

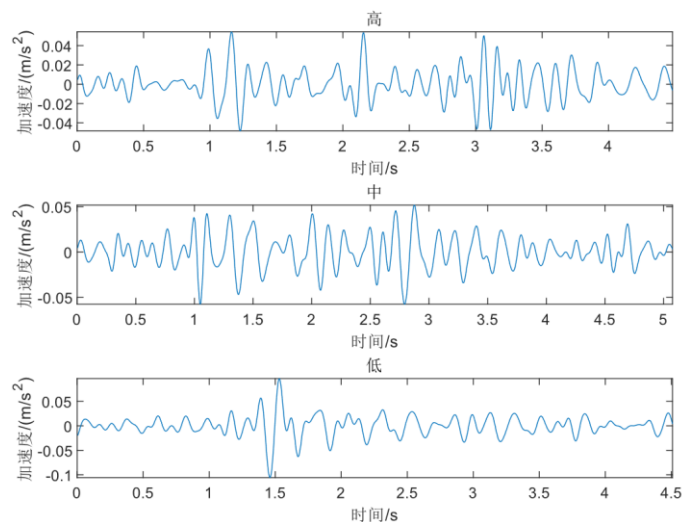


图 4-11 三种交互意图强度的 IAT 波形

#### 1) 构建初始特征集合

根据 4.3.2 小节内容,针对基于震颤信号的分析,本节首先需要对信号进行特征提取,获得一个初始的多维特征空间。具体而言,可以提取时域特征如均值、方差、四分位值、峰度峰值等,反映信号在时间轴上的统计特性。此外还可以提取频域特征比如不同频带下的功率谱密度,表示信号在频率分布上的信息。偏度、峰度等高阶统计量也能反映信号的非线性动态。主频率直接反映震颤的频率形态。所有特征构成了初始的 14 维特征向量空间,它为后续构建新特征提供了基础,14 维特征向量见表 4-5。

在表 4-5 中,从震颤信号的分布中,提取了 6 种时域特征、3 种频域特征和 3 种线性随机过程的特征。时域特征如均值、标准差等能反映信号幅值分布的集中趋势和离散程度,峰值特征反映信号的极大极小值,均方根反映信号能量大小。频域特征通过功率谱分析信号频率组成,主频率反映信号主要成分,随机过程特征如偏度和峰度描述信号概率分布的对称性和尖锐程度,线性对称性反映高阶统计量之间的相关性。综上,这组特征集合提供了丰富的统计学和信号处理角度的信息,能够可靠地刻画复杂信号的时域、频域和概率分布等关键属性。在此基础上,本小节可以进行更高层的特征选择、转换与建模,以实现特定的理解、诊断与预测任务。

表 4-5 构建特征集合

特征类别	特征名称
时域特征	均值(Mean)
	均方根(RMS)
	四分位点(Q1,Q2,Q3)

	标准差(SD)
	峰值(PV)
	峰-峰值(PPV)
频域特征	4-6 Hz 功率谱密度(LPSD)
	6-12 Hz 功率谱密度(HPSD)
	主频率 (FT)
线性随机过程特征	偏度(M3)
	峰度(M4)
	线性对称性(R)

## 2) 标准化处理

考虑到直接利用提取得到的 14 维原始时域和频域特征对复杂交互模式进行识别会带来一定的限制,本研究对特征进行适当预处理以消除特征值数量级上由量纲单位带来的差异。具体来说,可以对所有特征应用 Z-Score 标准化方法,使每个特征值分布映射到均值为 0 方差为 1 的统一高斯分布空间。Z-Score 标准化公式见公式 4-8。

$$Z = \frac{x - \mu}{\sigma} \quad (4-8)$$

在公式 4-8 中,  $Z$  表示标准化后的值,  $x$  是原始数据点的值,  $\mu$  是样本数据的均值,  $\sigma$  是样本数据的标准差。通过这个公式,原始数据点  $x$  被转换为相对于均值  $\mu$  的偏差,然后除以标准差  $\sigma$  进行标准化。这样处理后, Z-score 表示数据点在原始分布中的位置,负值表示在均值左侧,正值表示在均值右侧,而数值的大小表示与均值的偏离程度。

通过这种线性转换,不同特征之间的绝对差值会被相对关系比例替代。这样在新的映射空间内,所有特征都可比较,都会对后续的模式识别产生均衡影响。标准化操作还使信号的稳定统计特性被保留,有助于更稳健地反映交互动态。采用这类量纲归一化的预处理技术,是使多个异质特征协同支持非线性判别模型(如 SVM 与随机森林等)训练的常用和有效方法。其结果是不同特征可比,模型参数更稳定,从而提高交互意图分类的整体性能。

本节手动提取出的特征涵盖时域、频域以及线性随机过程,但直接用于后续机器学习方法分类前,有必要对特征空间进行适当筛选与缩减。一方面,当前特征集合存在冗余信息,使用全部特征不仅难以提高模型性能,反而会引入过拟合风险。另一方面,高维特征在实际应用中也会大大增加计算量。基于对上述 14 维震颤信号初始特征的分析,本文使用斯皮尔曼等级相关系数进行特征选择,去除强相关特征,优化特征子集的规模与表达能力。

## 3) 相关性检验

在进行 Z-Score 标准化消除特征量纲后,用斯皮尔曼等级相关系数评估各特征之间的线性依赖性,该系数通过综合两个变量值之间的单调趋势,判断它们的统计相关程度。

具体来说,可以求取标准化后特征矩阵中任意两列之间的斯皮尔曼系数,根据系数的绝对值,选择互相关性较弱的特征构成最后使用的子集。这通过相关性过滤可以去除冗余特征,使得保留特征对目标交互意图的条件独立性更强,避免多重共线性问题。采用非冗余、独立性较好的低维特征子集,可以使得基于决策树、贝叶斯网络等机器学习模型实现对复杂交互行为的更稳定、准确判别与分类。该特征选择方法计算高效,且通过阈值设置可以平衡特征数目与预测性能。

通过先后采用 Z-Score 标准化和斯皮尔曼等级相关系数方法处理原始特征集合中的 14 个特征,获得了一个经过量纲归一化和相关性筛选的优化特征子集。该子集在消除了特征量级差异的基础上,进一步剔除了互相关性高的冗余特征。其中,标准化通过线性映射实现了不同特征在数量级分布上的一致化处理,有利于它们对后续模式识别任务的均衡作用。而斯皮尔曼系数过滤则依据变量间单调函数关系的紧密程度,移除了相关性高的重复特征。因此,保留的特征子集在表达复杂交互动态信息方面可互补。并在支持向量机、K 近邻法等机器学习模型判别交互意图时,发挥更为独立的条件概率作用。这样,构建于该优化特征上的分类器可以获得更高的判别稳定性、准确率与泛化能力。

#### 4) 机器学习模型

在进行了针对原始特征集合选择与转换优化后,得到了一个无量纲且低相关的特征子集。基于该优化特征,本文选择机器学习模型去识别表示不同意图的用户行为。具体来说,将研究和对比支持向量机(SVM)、K 近邻(KNN)、决策树与朴素贝叶斯分类器在使用该特征子集条件下的准确度和性能。

SVM 通过寻找高维空间的最优决策边界,可以处理复杂的交互特征。KNN 通过实例示例直观地判别交互类别。而决策树与朴素贝叶斯可以从特征匹配中学习可解释的交互识别规则。设计实验评估这些模型的准确率等性能,最终选择合适的机器学习算法,使用优化特征识别不同交互意图和交互意图强弱的行为,进而达到获取用户交互意图的目的。

#### 5) 结果评估指标

本文采用了准确率、召回率、F1 分数指标对分类模型的结果进行评估。其中,准确率直接反映了分类器将测试样本正确划分到真实类别的比例。F1 得分考虑了模型对每个类别的精确率与召回率的调和平均,评价了分类的精确度与覆盖面。而召回率特别强调了分类器对某个类别内部样本的正确识别比例。这三种指标从不同侧面反映了多类分类的性能好坏:准确率关注整体判别效果、F1 得分关注模型对每个类别区分特性的掌握程度、召回率对某特定类别内部样本覆盖面的成效。将它们一并采用给出全面的性能描述,并通过与其他基准模型的比较确定实验中分类器的效果提升与独特优势。采用这种多指标联合评估的方法,可以从多角度判断分类模型的综合效果。

本文采用重复 5 折交叉验证进行性能评估，令机器学习分类方法得到更稳定和可靠的精度指标。具体来说，首先将全样本集随机分成 5 个子集，每次选择 1 个子集作为测试集，余下 4 个子集合并形成训练集。在该训练集上训练模型，计算其在对应的测试集上的精确度指标。如此循环 5 次，每份子集分别作为一次测试集，最终获得 5 个交叉验证结果。对该 5 个结果求均值可以得到该次划分情况下的模型泛化能力评分。通过使用重复交叉验证，不同数据划分和模型训练情况下对精度影响的变化都有权衡，使得最终的泛化性能评估更加稳定准确，可以对分类器的效果进行更加可靠的衡量。

#### 4.3.4 实验分析

本小节旨在通过震颤信号特征识别用户行为中是否蕴含交互意图及识别交互意图的强弱，进而建立更自然的空中手势交互系统。本实验的实验目的为探究通过震颤信号识别用户行为是否蕴含交互意图及蕴含交互意图的强弱，并以此针对空中手势交互中存在的 Midas Touch 问题研究一种解决方法。

##### 1) 参与者与实验设备

实验征集 9 名用户，均为所在学校的大学生，包括 4 名女性和 3 名男性，年龄分布在 20 岁到 23 岁之间( $M = 20.55, SD = 1.013$ )，惯用手均为右手，均无吸烟习惯和手部脑部疾病历史。48 小时内有 1 位同学有饮酒行为，3 位同学有喝咖啡的行为。有两位同学近期有服药经历，分别食用维生素及感冒药治疗感冒、食用阿莫西林治疗喉咙痛。

实验设备为惠普投影仪，投影尺寸为 100 寸，采集装置为 mpu9250 的惯性传感器，采样率为 200 Hz。



图 4-12 实验环境与交互设备图

##### 2) 实验任务与实验设计

在本实验中，参与者需要根据提示控制光标移动到指定区域，并按照提示进行数据采集。本次实验要求参与者右手佩戴数据采集装置执行交互任务，每名参与者需要参与 3 个目标区宽度的交互任务，每个任务有 3 个实验组块，参与者每个实验组块中需要执行 10 轮。目标区宽度的 3 个水平采用 3×3 拉丁方在不同的参与者间进行均衡。



实验变量：目标区宽度(30pt、90pt、180pt)。

实验类型：组内设计。

实验过程中的提示分为视觉提示和听觉提示。视觉提示为交互区域高亮，提示用户需要将光标移动/停留在高亮区域内。听力提示为交互过程中的叮咚声，叮咚声提示用户要开始移动光标。

本实验根据 Fitts 定律模型<sup>[79]</sup>构建交互任务，根据公式 4-9 中的难度指数 (ID)，给了固定距离  $D$  下三种不同难度系数。

$$ID = \log_2 \left( \frac{D}{W} + 1 \right) \quad (4-9)$$

公式 4-9 为设计实验时研究人员应该使用难度指数 (ID) 的 Shannon 公式，其中  $D$  表示两个交互目标中线之间的距离， $W$  表示目标区的宽度。在本实验中，实验难度指数分别为 5.67、4.14、3.22，认为难度指数越高的交互任务用户需要做出的努力越大并且交互意图越强烈。

在通过实验得到用户执行任务过程中的指尖六轴数据后，对数据进行预处理操作，删除一些异常数据值。根据 4.3.2 小节中描述，提取不同行为下的震颤特征，进行数据统计与分析。本实验确保所有参与者信息的隐私受到保护，遵守人体研究的伦理准则。

本次实验探究用户含有交互意图的行为与震颤特征之间的联系。本小节首先对数据进行了预处理，根据 4.3.1 小节消除高频噪声信号以及低频的行为信号，之后进行数据筛选，删掉传输过程中出现错误的的数据。之后根据正常人震颤的线性随机过程特性，分别提取偏度 (M3)、峰度 (M4) 特征、线性对称性 (R) 特征以及 4-6 Hz 功率谱密度 (LPSD)、6-12 Hz 功率谱密度 (HPSD) 特征进行分析，并提取构造的特征，使用支持向量机、k 最近邻、决策树、朴素贝叶斯分别对三种难度下运动性震颤和三种类型的震颤进行分类，计算其准确程度。

### 3) 震颤信号收集

在实验中，使用惯性传感器收集参与者指尖加速度、陀螺仪六轴数据，在交互任务中收集用户的 RT、IAT、PAT 信息，如图 4-13，在区域 1 记录 7s 包含休息性震颤信号 (RT) 的指尖数据。在区域 2 记录包含蕴含交互意图动作性震颤信号 (IAT) 的指尖数据。在区域 3 记录包含无意图姿势性震颤信号 (PAT) 的指尖数据。实验中探究不同交互意图强度的 IAT，则是从任务中的区域 2 中收集。

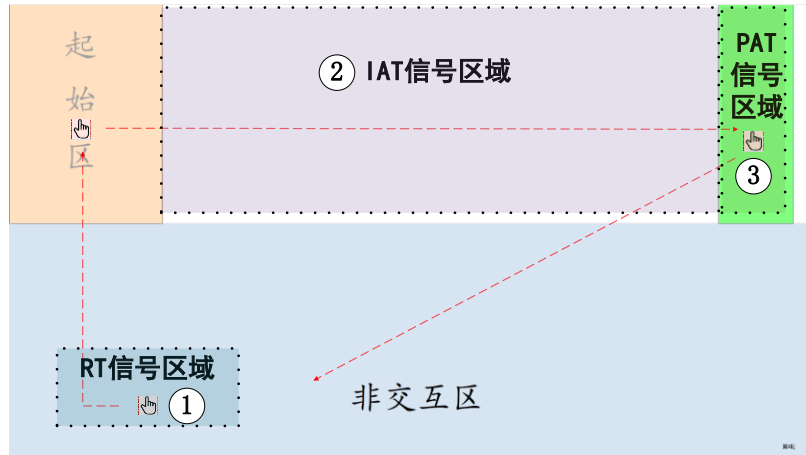


图 4-13 交互任务中收集用户信息区域

在本实验中，交互实验中采集用户的 RT、IAT、PAT 信息，用户状态分别为将手放在桌子上静止、在空中将手从一端旋转到另一端、在空中保持姿势静止。

#### 4) 实验过程

参与者在实验人员的帮助下佩戴好采集设备并熟悉了解任务流程，帮助参与者能够表示熟练交互系统和任务流程后结束，明确告知用户所有的声音提示都提示开始运动，所有的高亮提示都提示停留高亮区域。正式实验阶段，用户任务目标为从起始区出发移动到目标区，一次交互任务分为了 5 个阶段：1) 用户将手静止在桌面上，保持 7s 静止；2) 系统发出声音提示，用户将手抬起并将光标移动到起始区；3) 移动到起始区后，系统发出声音提示，用户将光标移动到目标区；4) 用户保持手部静止 7s，系统高亮目标区提示；5) 采集完成后，系统发出声音提示，用户将手移动到非交互区。针对交互过程中的 5 个阶段，其交互过程中关键位置如图 4-14 所示。

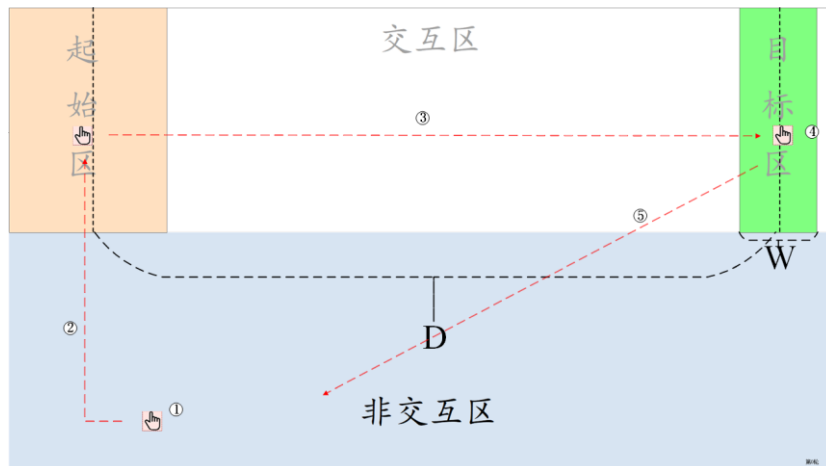


图 4-14 实验交互界面

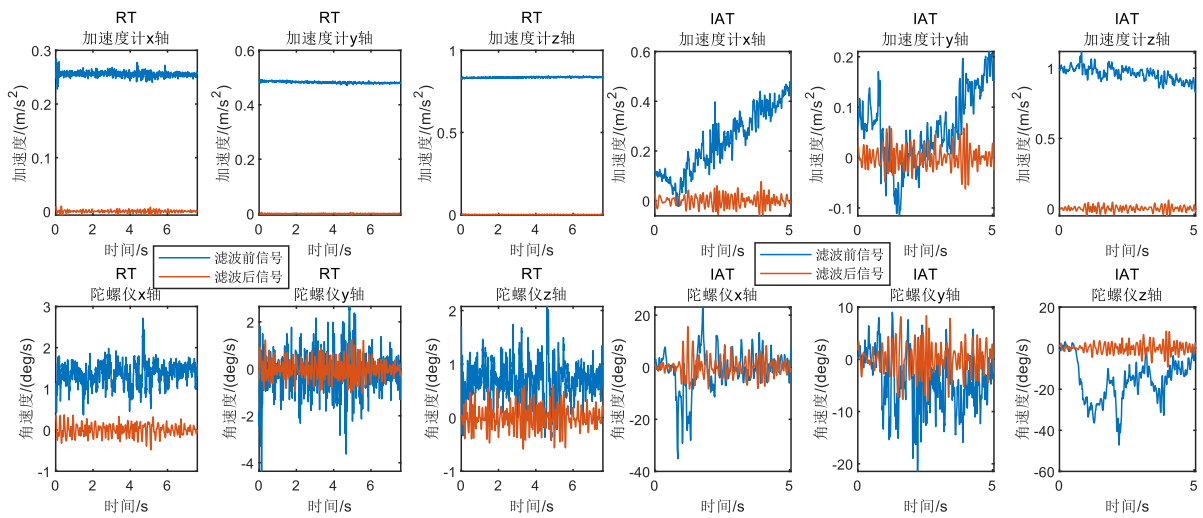
#### 5) 实验结果

通过实验，本研究获得了用户指尖加速度和角速度六轴原始数据，通过数据预处理，去除低频运动影响以及高频噪声，得到 RT、IAT、PAT 震颤信号。其中 IAT 为具有交互

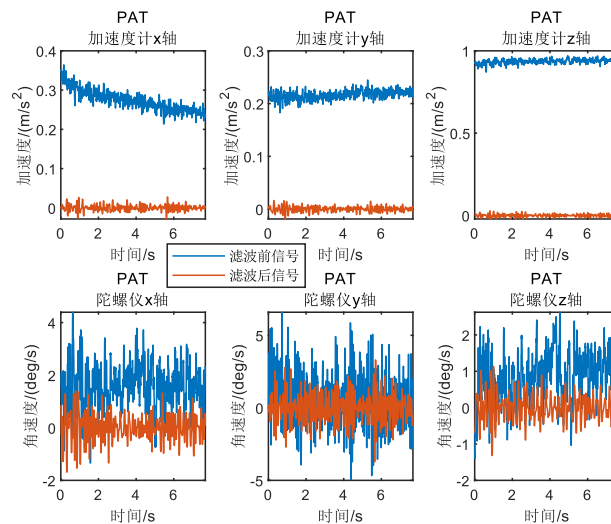


意图的运动性震颤，而其他两种无交互意图。其中，IAT 包含三种难度下即强、中、低交互意图的运动性震颤信号，本研究首先利用震颤特征对 RT、IAT、PAT 进行区分，之后探究不同交互意图是否影响三种信号的识别，进而为空中手势交互中 Midas Touch 问题研究一种解决方法。

首先对数据进行数据预处理，保留运动过程中的震颤信号，去除其他噪声。数据使用 10 阶巴特沃斯高通滤波器( $f_1 > 4 \text{ Hz}$ )进行带通滤波，用于消除手部动作的低频信号，再用 10 阶巴特沃斯低通滤波器( $f_2 < 12 \text{ Hz}$ )将高频噪声过滤掉，保留 RT 信号( $4 \text{ Hz} - 6 \text{ Hz}$ )与动作性震颤信号( $4 \text{ Hz} - 12 \text{ Hz}$ )。根据交互实验中标记与幅值检测，删除无效的数据，RT、IAT、PAT 滤波前后信号见图 4-15。



a) RT 滤波前后六轴信号    b) IAT 滤波前后六轴信号



c) PAT 滤波前后六轴信号

图 4-15 RT、IAT、PAT 滤波前后对比图

在对采集的手部和手指运动数据进行预处理后，得到 RT、IAT、PAT 的信号，其中 IAT 是用户含有交互意图的震颤信号，本实验接下来探究是否可以通过分析特征的差异，准确辨别出存在明确交互意图的 IAT 信号和无交互意图的 RT、IAT 信号。这样可为后续通过增强对复杂交互意图自身的识别，解决现有手势交互系统中存在的意图判断误操作问题提供可能性。

通过 4.3.1 小节所述，提取六轴信息的 m3、m4 特征，以实验中 block 为单位，以偏度为横轴，峰度为纵轴绘制分布图，见图 4-16。

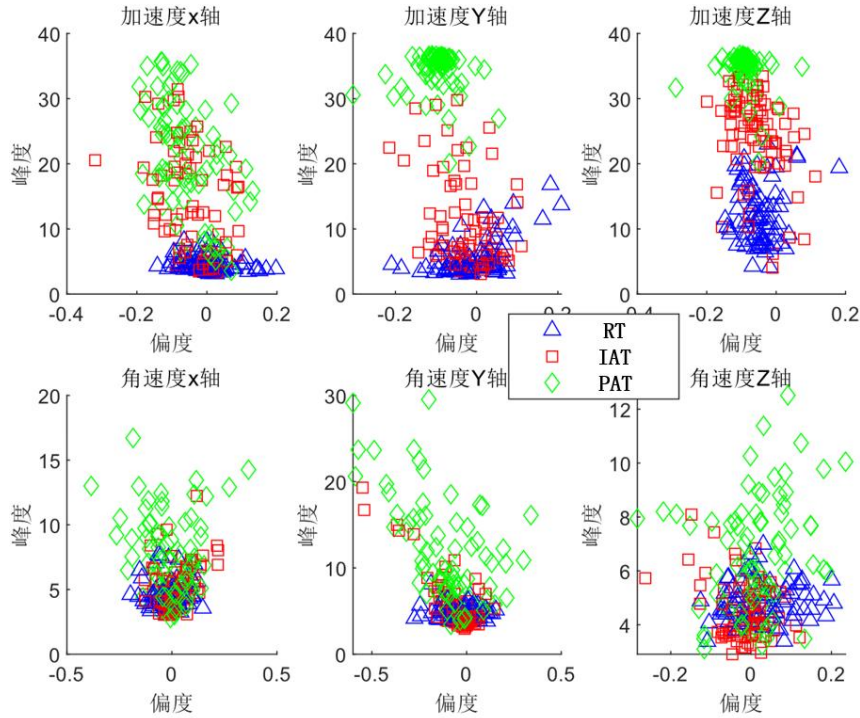


图 4-16 不同震颤类别的峰度偏度分布图

在图 4-16 中，三种震颤信号的峰度在加速度 Y 轴与 Z 轴表现出差异显著，通过 Wilcoxon 秩和检验中 RT 与 PAT 具有显著差异( $Z = -9.12, p < 0.01$ )，RT 与 IAT 具有显著差异( $Z = -10.9, p < 0.01$ )，在六轴数据中，RT 的峰度均值高于 PAT 和 IAT，PAT 和 IAT 在加速度 Z 轴差异显著( $Z = -10.2, p < 0.01$ )。峰度表示为信号分布的尖峭度，通过图 4-16 可以得出在加速度 Z 轴与 Y 轴上，RT 峰值最高，而 PAT 的峰值次之，IAT 信号峰值最低。

综上，仅通过峰度偏度分布，可以通过加速度 Z 轴与 Y 轴数据区分 RT、IAT、PAT。接下来本实验分析震颤的时间反演对称特征 R 进行进一步探究。

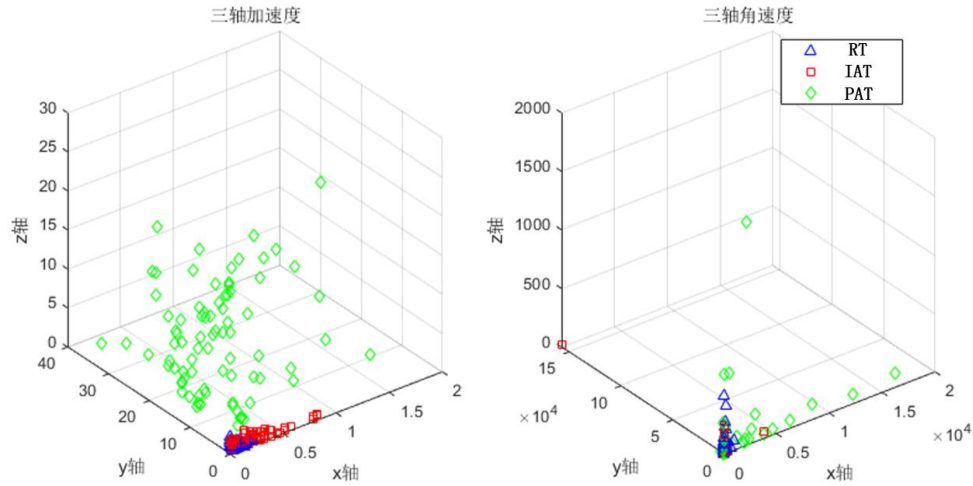


图 4-17 不同震颤类别的三轴分布图

在图 4-17 中，在加速度三轴构成的立体空间中，RT、IAT、PAT 分布出现分层，但同样无法通过该特征区分不同难度下的 IAT。为了进一步探究能否通过 LPSD 和 HPSD 进行区分，本实验通过提取六轴信号的 LPSD 和 HPSD，本实验以 LPSD 为横轴、以 HPSD 为纵轴，绘制 L-HPSD 分布图，如图 4-18。

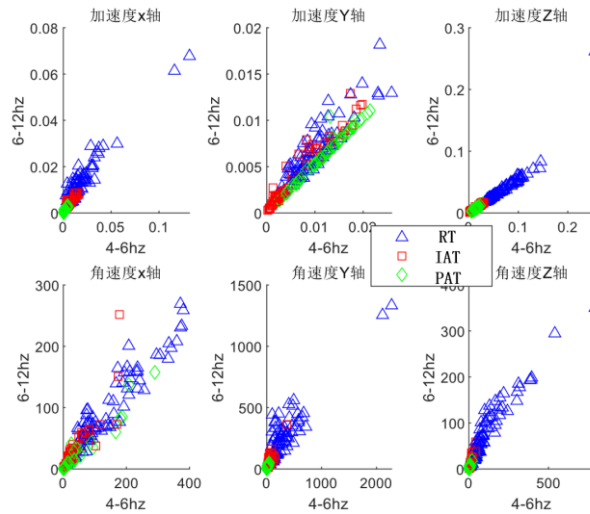


图 4-18 不同震颤类别的 LPSD-HPSD 分布图

在图 4-18 中，IAT 的 LPSD-HPSD 均值大于 RT 震颤和 PAT，这表明具有交互意图的震颤行为信号的 LPSD-HPSD 值均高于无交互意图的震颤行为。

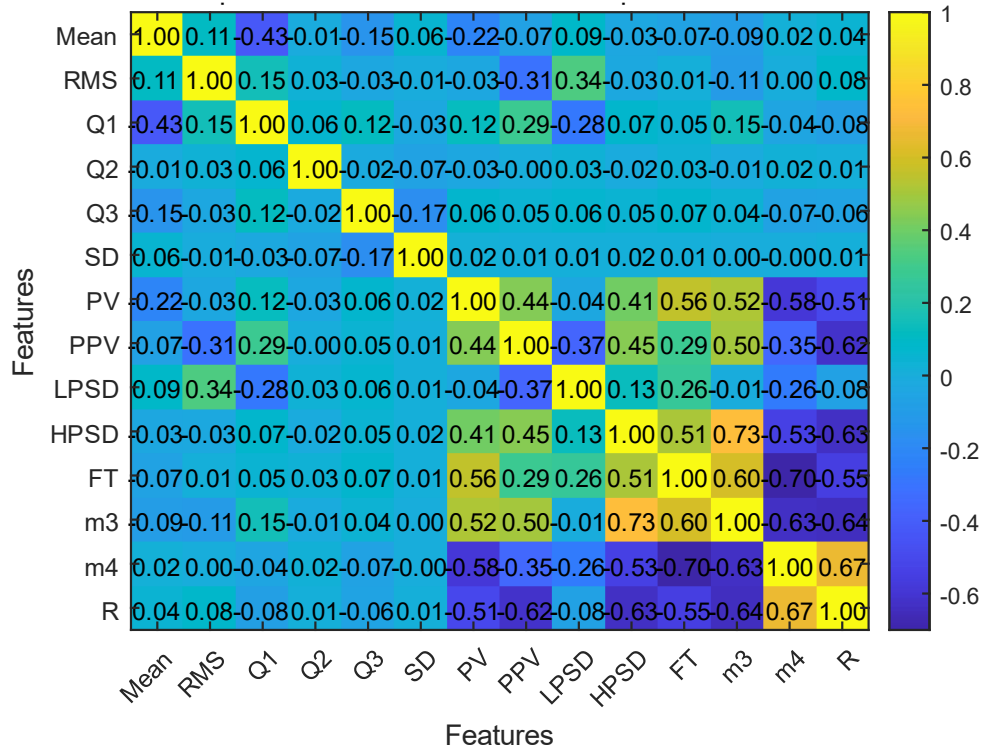


图 4-19 相关系数图

通过上述分析可知,仅通过简单的特征组合或单独特征,能够区分出 RT、IAT、PAT,本实验结合 4.3.3 节构建的特征集合,分别从六轴中提取特征,之后进行标准化处理,消除不同特征之间的尺度差异。之后针对特征集合求特征之间的斯皮尔曼等级相关系数,剔除相关系数较高的变量,特征之间相关系数结果如图 4-19。

在图 4-19 中,FT 和 m4、HPSP 和 m3 相关系数超过了 0.7,相关系数绝对值越接近 1,说明变量间的相关性越强,综合考虑,在研究之中剔除 HPSP 和 R 两个特征变量,之后使用 SVM、KNN、Decision Tree 和 Naive Bayes 进行分类。表 4-6 给出了 RT、IAT、PAT 分类的准确率、F1 分数和召回率。

表 4-6 区分 RT、IAT、PAT 结果

分类模型	Accuracy	F1 Score	Recall
SVM	0.305	0.289	0.315
KNN	0.837	0.829	0.849
Decision Tree	0.970	0.968	0.974
Naive Bayes	0.551	0.501	0.511

通过实验发现,在通过震颤特征对 RT、IAT、PAT 分类时,准确率最高为 97% (Decision Tree),通过震颤特征对 RT、IAT、PAT 分类具有较高的准确率,能够很好的识别三种行为震颤,接下来探究不同交互意图强度与区分 RT、IAT、PAT 是否有影响。

在实验中,根据不同目标区宽度划分为了3个难度的交互任务,难度系数越高的任务用户越专注,认为用户具有更高的交互意图,统计不同难度任务的执行时间如图4-20所示。

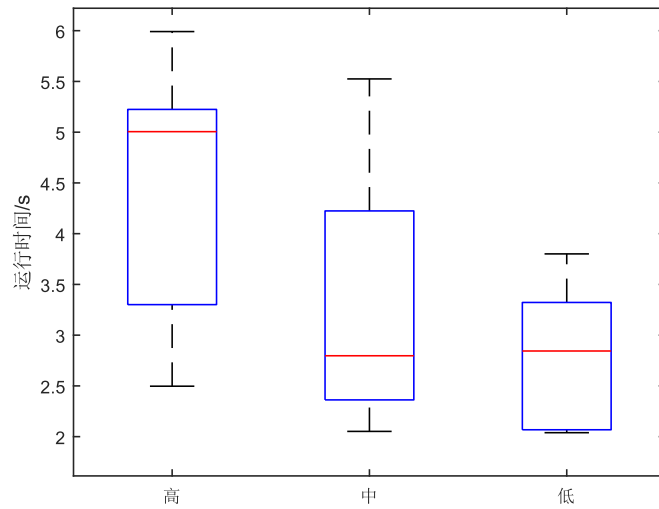


图 4-20 不同任务难度下平均用时箱线图

在图4-20中,图中分别为任务难度系数高、中和低下每名用户的平均执行时间分布情况,时间均值分别为4.36s、3.29s、2.82s。通过Lilliefors方法检验每种难度下的执行时间均符合正态分布,使用ANOVA进行显著性检验。实验结果展示,不同任务难度对实验用时有显著影响( $F(2,24) = 4.84, p < 0.05, \eta^2 = 0.28$ )。组内测试的对比结果显示平均用时在难度高与难度中之间没有表现出显著性差异( $F(1,16) = 3.40, p = 0.08$ ),难度高与难度低之间则表现出显著差异( $F(1,16) = 11.22, p < 0.01, \eta^2 = 0.41$ )。

综上所述,任务难度越高则用户所需时间均值越长,这表明不同难度的交互任务对执行时间有影响,在本实验中,认为用户在困难的任务中具有更加强烈的交互意图。为了进一步探究不同交互意图强度对于通过震颤区分用户行为是否有不同的表现,本研究基于实验设计中的三种不同强度交互任务,利用决策树进行区分,计算其准确率、F1分数及召回率,所得结果如图4-21所示。

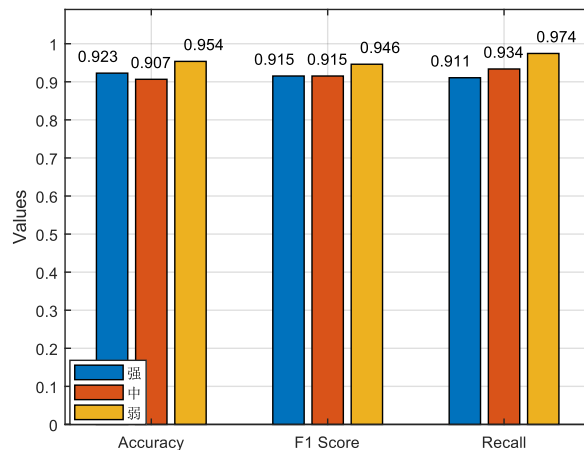


图 4-21 不同交互意图强度下区分 IAT、RT、PAT 结果

在图 4-21 不同强度交互意图下通过震颤识别用户行为中，准确率、F1 分数、召回率均超过了 90%，其中低交互意图相较于其他交互意图分类结果表现最好，但交互意图强弱与分类准确率之间并未表现出正相关或负相关。

综上所述，通过震颤信号的特征，区分 IAT 与无交互意图震颤信号具有较高的准确率 97%，而对于不同交互意图强度对于区分 IAT 与无交互意图震颤信号中，发现低意图强度下分类效果最好。因此，本文认为通过用户行为能够区分用户有无交互意图，这为设计更自然的人机交互方法提供了一种可行的解决方案。

#### 4.4 本章小结

本章旨在得到用户的手部姿态后，利用手部姿态进行手势交互，并完善手势交互中的相关问题。首先，利用以用户为中心的方法设计实验，实验过程中采用大声思考的方式获取用户对手势的设计想法，最终建立了基于姿态的手势集合。这可以避免现有手势集合无法满足本研究的平面手势交互需求和减轻设计者主观偏见的问题。其次，针对用户设计的手势集合，研究了一种坐标映射方法和手势识别方法。坐标映射方法为用户反馈的手掌偏航转动控制光标时转动角度范围小的问题提供了一种解决方法。手势识别方法则通过决策树利用手指关节特征进行分类。最后，本文分析了健康人群肌肉震颤信号与交互意图之间的关系，设计实验采集不同行为下的震颤特征，并利用多种分类器进行分类，区分 RT、IAT、PAT 时，决策树分类模型达到了 97% 的准确率，在不同交互意图强度下区分 IAT 时，准确率也均超过了 90%，这表明通过震颤信号识别交互意图具备可行性，进而通过识别用户行为是否具有交互意图时，系统选择是否进行功能映射，这为空中手势交互系统的误触与误识别问题提供了一种解决思路。

## 第 5 章 iGestureGlove 原型系统设计与实现

本章研究了一种模块化数据手套及空中手势交互原型系统 iGestureGlove。在第四章中,研究了一种基于姿态的手势集合及实现方法,探究了震颤信息与用户行为意图之间的联系,因此本章基于前文的研究,设计并实现了一个空中手势交互原型系统 iGestureGlove。

本章首先介绍了 iGestureGlove 原型系统的总体设计方案,包含系统硬件和软件两部分,在系统硬件部分中构建了一种模块化的数据手套,在软件部分介绍了本原型系统的软件结构。其次对本原型系统的实现进行进一步的介绍和展示,并介绍了本原型系统的应用场景。

### 5.1 iGestureGlove 的总体设计方案

iGestureGlove 手势交互原型系统从逻辑上分为硬件层、数据融合层、应用层三部分,iGestureGlove 系统框架如图 5-1 所示。

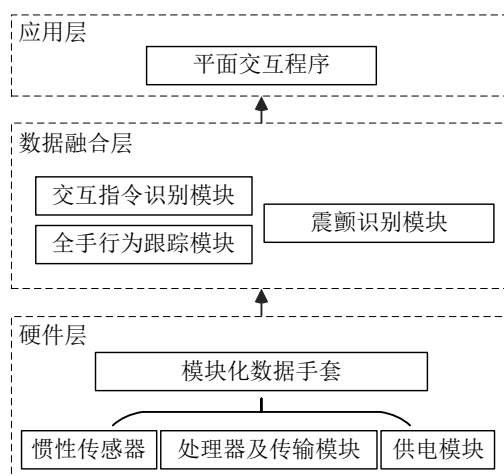


图 5-1 iGestureGlove 系统框架图

在图 5-1 中,硬件层作为系统的输入载体,所用模块化数据手套由惯性传感器、处理器、传输元件等设备组成,捕捉用户手部的详细运动数据,这为上层的手势跟踪和识别提供了基础数据。数据融合层由交互指令识别模块和震颤识别用户意图模块两部分组成,其中交互指令识别模块基于硬件层传感器的数据,采用数据融合算法,实时跟踪全手在三维空间中的行为,并基于所得姿态进行交互指令识别。识别用户意图模块则根据行为震颤特征识别用户行为是否包含交互意图。应用层的交互程序设计了多种平面交互任务,通过数据融合层提供的用户姿态及交互指令实现交互目的。

#### 5.1.1 原型系统硬件设计

iGestureGlove 原型系统的硬件部分为一种自研的模块化数据手套,其由惯性传感器、



无线传输模块、处理器、供电设备组成，原型硬件组成设计图如图 5-2 所示。模块化指将数据手套从物理层与逻辑层划分为若干个可扩充、可复用、抽象化的小模块。在本数据手套中，将惯性传感器作为最小单元，集成到单条柔性印制电路板(PCB, Printed Circuit Board)中，实现手指模块的设计，惯性传感器数目可以根据需求以及手指关节数目差异进行自主选择。其次将手指所附着的柔性 PCB 集成到手部 PCB 中，实现整手模块的设计，柔性 PCB 数目可以根据系统构型及功能需求自主选择。最后设计腕部 PCB，集成无线传输单元与处理器并与整手模块连接。

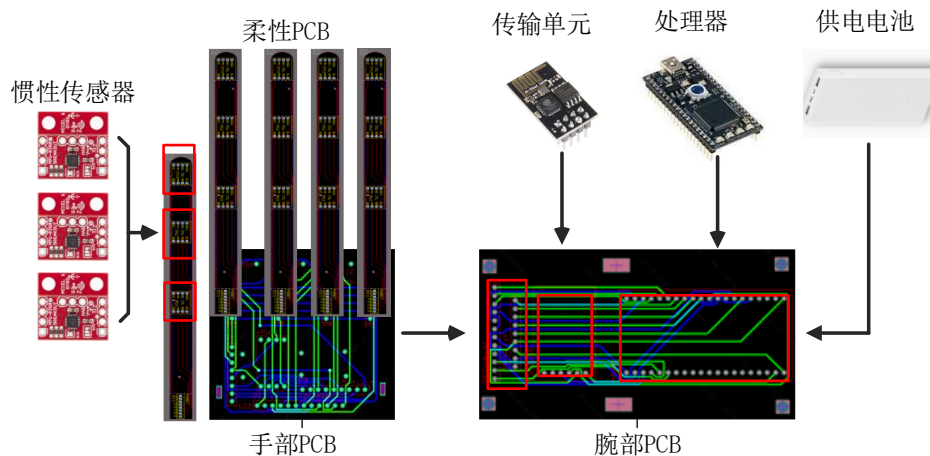


图 5-2 原型系统硬件组成设计图

在图 5-2 中，惯性传感器功能为采集附着骨节的运动信息，处理器则收集惯性传感器数据并整合为手部运动原始信息，进行数据预处理并输送到传输单元中，无线传输单元将手部原始信息通过 TCP/IP 网络串口发送到终端中。

iGestureGlove 原型系统的硬件部分所采用的器件清单如表 5-1 所示。

表 5-1 数据手套器件清单

名称	型号	功能
惯性传感器	SparkFun MPU-9250	获取手部运动信息
PCB	柔性板、FR-4 双层板	连通元件并提供模块化接口
处理器	mbed NXP LPC1768	读取并预处理传感器数据
传输单元	ATK-ESP8266	将处理后数据发送至终端
供电设备	18650 锂离子电池	为数据手套供电

在表 5-1 中，惯性传感器采用型号为 MPU-9250 的九轴传感器，这是一款融合了三轴加速度计、三轴陀螺仪和三轴磁力计的集成式惯性传感器模块，该传感器具有测量范围广、输出速率高、功耗低以及尺寸小等优点，通过附着骨节上提供手部运动信息。集成惯性传感器的柔性 PCB 和手部腕部 PCB 则是自主设计，分析手部构型以及惯性传感



器接口，提供一组可复用的单指以及全手连接装置。处理器采用的是 mbed 微控制器，具有低功耗和高性能计算能力等优点，通过 SPI 通信协议与惯性传感器进行数据传输，对原始数据进行预处理操作。传输单元采用型号为 ATK-ESP8266 的 Wi-Fi 通讯模块，具有性能稳定、接口丰富且成本较低的优点，通过 TCP 协议与终端实现数据通信，将手部运动信息从数据手套发送到计算机中。供电设备则采用由 18650 锂离子电池组成的便携充电设备中，该模块能够为数据采集装置提供稳定的 5V 电压，满足工作过程的能量需求。原型系统所用模块化数据手套具有低耦合、高复用、拓展性强、便携的特点，能满足不同交互环境的需求，实物图如图 5-3 所示。



图 5-3 iGestureGlove 模块化数据手套

### 5.1.2 原型系统软件结构

iGestureGlove 主要包括数据采集模块、数据处理模块及交互应用模块，其具体结构如图 5-4 所示。

iGestureGlove 原型系统结构中的数据采集层部署在微控制器中，所用开发语言为 C++ 语言，根据需求预定义柔性 PCB 及惯性传感器数目，使用 SPI 通信协议获取惯性传感器的角速度、加速度及磁场强度信息，并使用 TCP 通信协议通过网络串口传输到计算机中。数据处理层和交互应用层在计算机中实现，数据处理层中手部姿态求解算法为第三章内容，震颤特征提取以及交互意图识别则使用第四章中模型识别，交互指令映射与交互手势识别则通过第四章的决策树模型实现，交互应用层所用开发语言为 C# 语言，开发平台为 Visual Studio，交互应用层通过虚拟串口获取数据处理层的数据，提供了展示手部姿态、执行交互任务的功能。

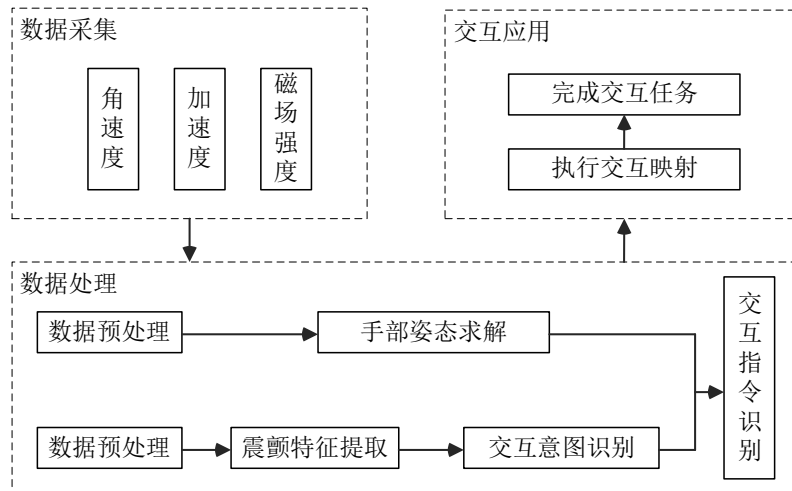


图 5-4 原型系统软件结构图

iGestureGlove 原型系统软件的开发环境如表 5-2 所示。

表 5-2 软件开发环境清单

开发环境	版本
操作系统	Windows 11, 64 bit
开发语言	C++, C#
开发工具	Keil Studio, Visual Studio 2019
依赖开发包	.NET FrameWork4.0

## 5.2 iGestureGlove 原型系统实现与应用

iGestureGlove 原型系统的软件界面实现如图 5-5、图 5-6 所示，主要由全手行为跟踪、交互意图识别和手势交互系统三部分组成。用户可以通过该原型系统得出手部姿态的精确信息，通过虚拟手的视觉反馈观察手部状态。此外还可以得出当前行为的震颤信号，通过用户行为震颤特征判定用户行为是否包含交互意图。最后，用户可以执行交互模块中预设的交互任务，实现自然的手势交互。

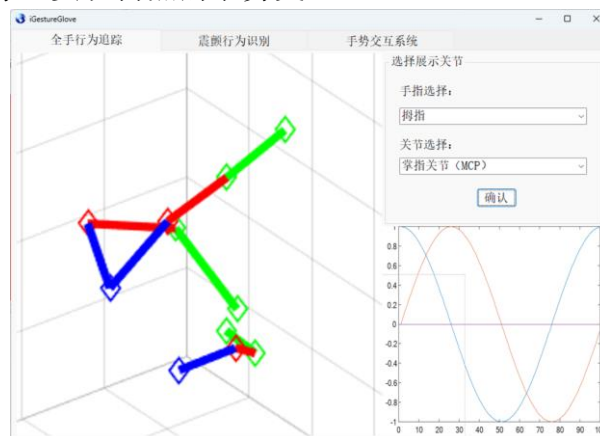


图 5-5 虚拟手及关节姿态界面

在图 5-5 中, 该部分可将手部姿态以虚拟手的形式展示在界面中, 通过佩戴数据手套执行手部动作, 虚拟手同步展示在界面左边区域, 同时还可以令用户自主选择关节查看关节姿态。iGestureGlove 原型系统通过提取震颤特征以实现用户意图检测, 选用了第四章中准确率最高的决策树分类模型进行识别, 判断当前阶段用户行为是否为包含交互意图。

iGestureGlove 原型系统提供的交互模块则允许用户通过手势完成交互任务, 系统实现的交互指令包含移动、选择、抓取、旋转和缩放五类交互功能, 使用第四章中以用户为中心的手势集合, 通过映射函数实现对光标的平移控制, 通过决策树模型实现手势的分类, 为手势交互提供基础支持。

基于上述研究, 本系统在交互模块中设计了四种交互任务, 分别是: 从多个对象中选择指定对象、移动对象到指定区域、旋转对象到指定角度和缩放对象到目标大小。交互任务界面如图 5-6 所示。



图 5-6 平面交互系统界面

传统基于镜像的镜像疗法理论认为通过观察非患病肢体移动可以激活病侧脑区, 促进功能康复。但是这种方法存在自身运动感觉反馈不足的缺点, 患者难以准确判断肢体移动的细微变化。而 iGestureGlove 系统可以准确获取各手部关节的运动数据, 并通过虚拟手显示为用户提供视觉反馈, 如图 5-7 所示。



图 5-7 传统镜像反馈疗法<sup>[80]</sup>与交互系统反馈

相比传统镜像疗法，iGestureGlove 系统克服了反馈不足的问题，可以让患者更准确地感知自己的肢体运动。同时，虚拟手的显示也增强了视觉引导作用，有助于激活患病侧脑区，促进运动功能改善，基于 iGestureGlove 系统的镜像疗法可以获得比传统镜像疗法更好的康复效果。

图 5-8 展示了 iGestureGlove 原型系统在大屏环境下的应用场景。用户戴着数据手套，可以在行走或移动时进行手势交互操作，而不受空间位置的限制。系统可以通过判断用户行为是否包含交互意图，从而实现流畅的手势交互。这展示了 iGestureGlove 允许用户突破固定桌面环境的限制，实现真正的移动互动体验。相比被动的手势识别交互，iGestureGlove 实现了基于意图的主动交互模式，大大提高了流动环境下的交互效率与用户体验。

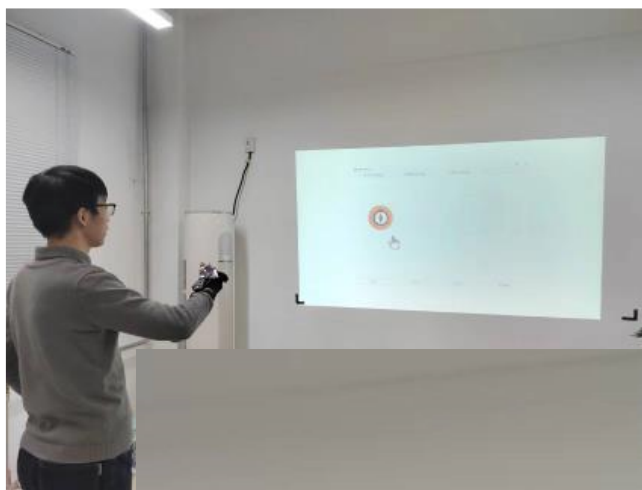


图 5-8 大屏环境下的应用场景

### 5.3 本章小结

本章围绕 iGestureGlove 原型系统设计与实现展开论述，首先探究了一种模块化数据手套的设计方案与实例。其次阐述了 iGestureGlove 原型系统的软件结构，分别包括部署在微处理器的数据采集层和部署在计算机端的数据处理层和交互层。最后展示了 iGestureGlove 原型系统的实现与应用。

## 第6章 总结与展望

### 6.1 总结

为了提高手势交互系统中手势识别的准确率,降低交互延迟,为用户提供一种自然流畅的交互体验,本文研究高可用单手空中手势交互技术,现将本文的工作做如下总结:

1) 研究了一种融合手部约束的全手行为跟踪算法,降低了可穿戴设备存在的错位误差。通过发掘人体手部结构与关节约束,将关节的生理约束转化为基于姿态描述的数值约束,并将其作为先验知识融入到姿态求解算法。本方法无需额外的设备支持并且可扩展到更多的约束维度,以提高姿态估计方法的性能。

2) 基于以用户为中心的方法设计了一套手势集合,该手势集合满足手势交互的基本功能。本文基于该手势集合给出了一种实现方法,通过长短距坐标映射模型,能够满足人手活动空间与虚拟平面空间的坐标映射,通过手部姿态的决策树模型识别交互指令。此外本文探究了震颤与交互意图联系,实验验证了通过震颤识别交互意图的可行性,为手势交互的误识别问题提供一种解决思路。

3) 设计并实现了一个空中交互的原型系统 iGestureGlove,研究了一种模块化数据手套的结构与实例。由于模块化的设计,本文的数据手套可以满足多种手部模型需求,具有较高的拓展性。本文设计的原型系统通过数据手套捕捉用户的运动信息,与实际应用场景相结合,实现自然手势交互。

### 6.2 展望

本文工作面向基于姿态的手部交互领域,对手部姿态跟踪和交互技术进行了一定的研究和讨论,但目前的工作仍存在一些不足,需要进一步地研究和完善:

1) 本文采用的手部约束方法主要将约束信息转化为基于姿态的范围约束,然而并未充分考虑手部关节的肌肉运动关系。举例而言,掌指关节在手部运动中涉及两个自由度,而这两个自由度之间存在相互制约的关系,两个自由度的姿态值无法同时达到最大或最小值。因此,可以根据手部肌肉运动原理,对手部约束进行精细化建模,以进一步提高姿态估计方法的准确性。

2) 本文所研究的五种手势类型满足手势交互的基本功能,但更复杂的综合交互场景需要更多的功能实现,有必要增加研究的手势类型,以丰富手势的多样性。通过引入更多手势类型,能够更全面地捕捉用户的意图和动作,从而提升交互系统的灵活性和适用性。

3) 本文着重探讨了通过震颤特征识别实现交互意图的可行性,未来深入挖掘震颤特征是如何与交互意图关联的,为了更全面地理解和解释震颤特征与交互意图之间的关

系，同时研究进一步拓展与细分用户行为，定位影响交互意图的震颤特征，减少手势交互中的误触和误识别问题。

4) 将交互意图纳入交互设计和算法优化的考虑中，有助于建立更智能、用户友好的交互系统。未来的研究方向可以包括对不同交互场景下交互意图的更精细化分析，以提高系统对用户意图的理解和识别能力，从而进一步减少交互过程中的误触和误识别。

## 参考文献

- [1] Khan A A, Newn J, Bailey J, et al. Integrating Gaze and Speech for Enabling Implicit Interactions[C]//Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022: 1-14.
- [2] Bai H, Sasikumar P, Yang J, et al. A user study on mixed reality remote collaboration with eye gaze and hand gesture sharing[C]//Proceedings of the 2020 CHI conference on human factors in computing systems. 2020: 1-13.
- [3] Hendrickson W B. The three lives of Frank H. Hall[J]. Journal of the Illinois State Historical Society (1908-1984), 1956, 49(3): 271-293.
- [4] Loguidice B, Barton M. Vintage Game Consoles: an inside look at Apple, Atari, Commodore, Nintendo, and the greatest gaming platforms of all time[M]. CRC Press, 2014.
- [5] Korayem M H, Madihi M A, Vahidifar V. Controlling surgical robot arm using leap motion controller with Kalman filter[J]. Measurement, 2021, 178: 109372.
- [6] Kang P, Li J, Jiang S, et al. Reduce system redundancy and optimize sensor disposition for EMG–IMU multimodal fusion human–machine interfaces with XAI[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 72: 1-9.
- [7] Hassan W, Raza A, Abdullah M, et al. HapWheel: Bringing In-Car Controls to Driver’s Fingertips by Embedding Ubiquitous Haptic Displays into a Steering Wheel[J]. IEEE transactions on intelligent transportation systems, 2022, 23(10): 18526-18534.
- [8] Hisham B, Hamouda A. Supervised learning classifiers for Arabic gestures recognition using Kinect V2[J]. SN Applied Sciences, 2019, 1: 1-21.
- [9] Zhu Y, Lu W, Gan W, et al. A contactless method to measure real-time finger motion using depth-based pose estimation[J]. Computers in Biology and Medicine, 2021, 131: 104282.
- [10] Antico M, Balletti N, Laudato G, et al. Postural control assessment via Microsoft Azure Kinect DK: An evaluation study[J]. Computer Methods and Programs in Biomedicine, 2021, 209: 106324.
- [11] Guerra-Segura E, Ortega-Pérez A, Travieso C M. In-air signature verification system using leap motion[J]. Expert Systems with Applications, 2021, 165: 113797.
- [12] Cai Y, Ge L, Cai J, et al. 3D hand pose estimation using synthetic data and weakly labeled RGB images[J]. IEEE transactions on pattern analysis and machine intelligence, 2020, 43(11): 3739-3753.
- [13] Avola D, Cinque L, Fagioli A, et al. 3D hand pose and shape estimation from RGB images for keypoint-based hand gesture recognition[J]. Pattern Recognition, 2022, 129: 108762.

- 
- [14] Li Y, Wang X, Liu W, et al. Deep attention network for joint hand gesture localization and recognition using static RGB-D images[J]. *Information Sciences*, 2018, 441: 66-78.
- [15] Pinto R F, Borges C D B, Almeida A M A, et al. Static hand gesture recognition based on convolutional neural networks[J]. *Journal of Electrical and Computer Engineering*, 2019, 2019: 1-12.
- [16] Chen Z, Sun Y. Joint-wise 2D to 3D lifting for hand pose estimation from a single RGB image[J]. *Applied Intelligence*, 2023, 53(6): 6421-6431.
- [17] Streli P, Armani R, Cheng Y F, et al. HOOV: Hand Out-Of-View Tracking for Proprioceptive Interaction using Inertial Sensing[C]//*Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023: 1-16.
- [18] Bahl P, Padmanabhan V N. RADAR: An in-building RF-based user location and tracking system[C]//*Proceedings IEEE INFOCOM 2000. Conference on computer communications. Nineteenth annual joint conference of the IEEE computer and communications societies (Cat. No. 00CH37064)*. Ieee, 2000, 2: 775-784.
- [19] Pu Q, Gupta S, Gollakota S, et al. Whole-home gesture recognition using wireless signals[C]//*Proceedings of the 19th annual international conference on Mobile computing & networking*. 2013: 27-38.
- [20] Zhang J, Li Y, Xiong H, et al. HandGest: Hierarchical sensing for robust-in-the-air handwriting recognition with commodity WiFi devices[J]. *IEEE Internet of Things Journal*, 2022, 9(19): 19529-19544.
- [21] Han Z, Lu Z, Wen X, et al. CentiTrack: Towards centimeter-level passive gesture tracking with commodity WiFi[J]. *IEEE Internet of Things Journal*, 2023.
- [22] Zhang O, Srinivasan K. Mudra: User-friendly fine-grained gesture recognition using WiFi signals[C]//*Proceedings of the 12th International on Conference on emerging Networking EXperiments and Technologies*. 2016: 83-96.
- [23] Wang Y, Wu K, Ni L M. Wifall: Device-free fall detection by wireless networks[J]. *IEEE Transactions on Mobile Computing*, 2016, 16(2): 581-594.
- [24] Abuhoureyah F S, Wong Y C, Isira A S B M. WiFi-based human activity recognition through wall using deep learning[J]. *Engineering Applications of Artificial Intelligence*, 2024, 127: 107171.
- [25] Kortier H G, Sluiter V I, Roetenberg D, et al. Assessment of hand kinematics using inertial and magnetic sensors[J]. *Journal of neuroengineering and rehabilitation*, 2014, 11(1): 1-15.
- [26] Fang B, Sun F, Liu H, et al. 3D human gesture capturing and recognition by the IMMU-based data glove[J]. *Neurocomputing*, 2018, 277: 198-207.



- [27] Lin B S, Lee I J, Chiang P Y, et al. A modular data glove system for finger and hand motion capture based on inertial sensors[J]. Journal of medical and biological engineering, 2019, 39: 532-540.
- [28] Lin B S, Lee I J, Hsiao P C, et al. An assessment system for post-stroke manual dexterity using principal component analysis and logistic regression[J]. IEEE transactions on neural systems and rehabilitation engineering, 2019, 27(8): 1626-1634.
- [29] Chang H T, Chang J Y. Sensor glove based on novel inertial sensor fusion control algorithm for 3-D real-time hand gestures measurements[J]. IEEE Transactions on Industrial Electronics, 2019, 67(1): 658-666.
- [30] Jiang L, Xia H, Guo C. A model-based system for real-time articulated hand tracking using a simple data glove and a depth camera[J]. Sensors, 2019, 19(21): 4680.
- [31] Park G, Argyros A, Lee J, et al. 3D hand tracking in the presence of excessive motion blur[J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 26(5): 1891-1901.
- [32] 刘武. 基于多元信息融合的手指跟踪方法及其实现[D]. 西安. 西北工业大学,2019.
- [33] 李泽. 基于磁感应和信息融合的手势跟踪技术研究[D]. 西安. 西北工业大学,2019
- [34] 刘兴芮. 融合脑电信号的手部行为预测及跟踪技术研究[D]. 西安. 西北工业大学,2022
- [35] Lin J, Wu Y, Huang T S. Modeling the constraints of human hand motion[C]//Proceedings workshop on human motion. IEEE, 2000: 121-126.
- [36] Cerveri P, De Momi E, Lopomo N, et al. Finger kinematic modeling and real-time hand motion estimation[J]. Annals of biomedical engineering, 2007, 35: 1989-2002.
- [37] 温豪. 自主手部跟踪和交互技术研究[D]. 西安. 西北工业大学,2022
- [38] Li Y, Zhang L, Qiu Z, et al. NIMBLE: a non-rigid hand model with bones and muscles[J]. ACM Transactions on Graphics (TOG), 2022, 41(4): 1-16.
- [39] Cheng H, Yang L, Liu Z. Survey on 3D hand gesture recognition[J]. IEEE transactions on circuits and systems for video technology, 2015, 26(9): 1659-1673.
- [40] Feix T. Anthropomorphic hand optimization based on a latent space analysis[D]. Technische Universität Wien, 2011.
- [41] Cui J W, Li Z G, Du H, et al. Recognition of Upper Limb Action Intention Based on IMU[J]. Sensors, 2022, 22(5): 1954.
- [42] Tian Z, Ren M, Jiang Q, et al. Wi-Fi Gesture Recognition Technology Based on Time-Frequency Features. In Proceedings of the International Conference on Communications and Networking in China; Springer: Cham, Switzerland, 2019, 402–413.

- 
- [43] Ulhaq A, Yin X, He J, et al. On space-time filtering framework for matching human actions across different viewpoints[J]. IEEE Transactions on Image Processing, 2017, 27(3): 1230-1242.
- [44] Ji R. Research on basketball shooting action based on image feature extraction and machine learning[J]. IEEE Access, 2020, 8: 138743-138751.
- [45] Mishchenko Y, Kaya M, Ozbay E, et al. Developing a three-to six-state EEG-based brain-computer interface for a virtual robotic manipulator control[J]. IEEE transactions on biomedical engineering, 2018, 66(4): 977-987.
- [46] Zhong T, Li D, Wang J, et al. Fusion learning for sEMG recognition of multiple upper-limb rehabilitation movements[J]. Sensors, 2021, 21(16): 5385.
- [47] Ning N, Tang Y. Evaluation of an Information Flow Gain Algorithm for Microsensor Information Flow in Limber Motor Rehabilitation[J]. Complexity, 2021, 2021: 1-11.
- [48] Hussain G, Jabbar M S, Cho J D, et al. Indoor positioning system: A new approach based on lstm and two stage activity classification[J]. Electronics, 2019, 8(4): 375.
- [49] Sun Y, Huang R, Zheng J, et al. Design and speed-adaptive control of a powered geared five-bar prosthetic knee using bp neural network gait recognition[J]. Sensors, 2019, 19(21): 4662.
- [50] Strachan S, Murray-Smith R. Muscle tremor as an input mechanism[C]//Proc. UIST. 2004, 2: 482.
- [51] Meier P, Rohrmann K, Sandner M, et al. A novel methodology for magnetic hand motion tracking in human-machine interfaces[C]//2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2018: 372-378.
- [52] Pasku V, De Angelis A, De Angelis G, et al. Magnetic field analysis for 3-D positioning applications[J]. IEEE Transactions on Instrumentation and Measurement, 2017, 66(5): 935-943.
- [53] Santoni F, De Angelis A, Moschitta A, et al. MagIK: A hand-tracking magnetic positioning system based on a kinematic model of the hand[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-13.
- [54] Molchanov P, Yang X, Gupta S, et al. Online detection and classification of dynamic hand gestures with recurrent 3d convolutional neural network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 4207-4215.
- [55] Cihan Camgoz N, Hadfield S, Koller O, et al. Subunets: End-to-end hand shape and continuous sign language recognition[C]//Proceedings of the IEEE international conference on computer vision. 2017: 3056-3065.

- 
- [56] Cao C, Zhang Y, Wu Y, et al. Egocentric gesture recognition using recurrent 3d convolutional neural networks with spatiotemporal transformer modules[C]//Proceedings of the IEEE international conference on computer vision. 2017: 3763-3771.
- [57] De Smedt Q, Wannous H, Vandeborre J P. Heterogeneous hand gesture recognition using 3D dynamic skeletal data[J]. Computer Vision and Image Understanding, 2019, 181: 60-72.
- [58] Boulahia S Y, Anquetil E, Multon F, et al. Dynamic hand gesture recognition based on 3D pattern assembled trajectories[C]//2017 seventh international conference on image processing theory, tools and applications (IPTA). IEEE, 2017: 1-6.
- [59] Zhao D, Liu Y, Li G. Skeleton-based dynamic hand gesture recognition using 3d depth data[J]. Electronic Imaging, 2018, 30: 1-8.
- [60] Shuster M D. A survey of attitude representations[J]. Navigation, 1993, 8(9): 439-517.
- [61] Choi Y, Yoo K, Kang S J, et al. Development of a low-cost wearable sensing glove with multiple inertial sensors and a light and fast orientation estimation algorithm[J]. The Journal of Supercomputing, 2018, 74: 3639-3652.
- [62] Salchow-Hömmen C, Callies L, Laidig D, et al. A tangible solution for hand motion tracking in clinical applications[J]. Sensors, 2019, 19(1): 208.
- [63] Wikipedia. Hand. [DB/OL].<https://en.wikipedia.org/w/index.php?title=Hand&oldid=1054072982>, 2021-11-07
- [64] Cobos S, Ferre M, Uran M A S, et al. Efficient human hand kinematics for manipulation tasks[C]//2008 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2008: 2246-2251.
- [65] Valenti R G, Dryanovski I, Xiao J. A linear Kalman filter for MARG orientation estimation using the algebraic quaternion algorithm[J]. IEEE Transactions on Instrumentation and Measurement, 2015, 65(2): 467-481.
- [66] Wobbrock J O, Findlater L, Gergle D, et al. The aligned rank transform for nonparametric factorial analyses using only anova procedures[C]//Proceedings of the SIGCHI conference on human factors in computing systems. 2011: 143-146.
- [67] Hosseini M, Ihmels T, Chen Z, et al. Towards a Consensus Gesture Set: A Survey of Mid-Air Gestures in HCI for Maximized Agreement Across Domains[C]//Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 2023: 1-24.
- [68] Wobbrock J O, Morris M R, Wilson A D. User-defined gestures for surface computing[C]//Proceedings of the SIGCHI conference on human factors in computing systems. 2009: 1083-1092.

- 
- [69] Vatavu R D, Wobbrock J O. Formalizing agreement analysis for elicitation studies: new measures, significance test, and toolkit[C]//Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015: 1325-1334.
- [70] Poupyrev I, Billingham M, Weghorst S, et al. The go-go interaction technique: non-linear mapping for direct manipulation in VR[C]//Proceedings of the 9th annual ACM symposium on User interface software and technology. 1996: 79-80.
- [71] Jones S F. The Physiology of the Joints[J]. Physiotherapy, 1990, 76(1): 52-52.
- [72] Kandel, Eric R. Principles of neural science[J]. Oxford, 1983.
- [73] Elble R J. Physiologic and essential tremor[J]. Neurology, 1986, 36(2): 225-225.
- [74] Deuschl G, Bain P, Brin M, et al. Consensus statement of the movement disorder society on tremor[J]. Movement disorders, 1998, 13(S3): 2-23.
- [75] Elble R J, Randall J E. Motor-unit activity responsible for 8-to 12-Hz component of human physiological finger tremor[J]. Journal of Neurophysiology, 1976, 39(2): 370-383.
- [76] López-Blanco R, Velasco M A, Méndez-Guerrero A, et al. Essential tremor quantification based on the combined use of a smartphone and a smartwatch: The NetMD study[J]. Journal of Neuroscience Methods, 2018, 303: 95-102.
- [77] Timmer J, Häußler S, Lauk M, et al. Pathological tremors: Deterministic chaos or nonlinear stochastic oscillators?[J]. Chaos: An Interdisciplinary Journal of Nonlinear Science, 2000, 10(1): 278-288.
- [78] Edwards R, Beuter A. Using time domain characteristics to discriminate physiologic and parkinsonian tremors[J]. Journal of Clinical Neurophysiology, 2000, 17(1): 87-100.
- [79] Card S K, English W K, Burr B J. Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT[J]. Ergonomics, 1978, 21(8): 601-613.
- [80] Mannino R. Using mirror therapy to reduce pain and improve movement[J]. 2021.

## 致 谢

## 攻读硕士学位期间发表的学术论文和参加科研情况

攻读硕士期间发表的专利情况如下：

一种融合约束的手部姿态测量方法

申请号：CN202211185575.6

公开号：CN115761787A

公开日：2023.03.07

攻读硕士期间参与的软件著作权情况如下：

动态空中手势早期识别系统软件 V1.0

开发完成日期：2023.05.28

登记号：2023SR1493092

# 西北工业大学

## 学位论文知识产权声明书

本人完全了解学校有关保护知识产权的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属于西北工业大学。学校有权保留并向国家有关部门或机构送交论文的复印件和电子版。本人允许论文被查阅和借阅。学校可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。同时本人保证，毕业后结合学位论文研究课题再撰写的文章一律注明作者单位为西北工业大学。

本学位论文属于（在以下方框内打“√”）：

☐ 保密论文，保密期（     年     月     日至     年     月     日）。

☐ 公开论文。

学位论文作者签名：\_\_\_\_\_

年    月    日

指导教师签名：\_\_\_\_\_

年    月    日

---

# 西北工业大学

## 学位论文原创性声明

秉承学校严谨的学风和优良的科学道德，本人郑重声明：所呈交的学位论文，是本人在导师的指导下进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容和致谢的地方外，本论文不包含任何其他个人或集体已经公开发表或撰写过的研究成果，不包含本人或其他已申请学位或其他用途使用过的成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式表明。

本人学位论文与资料若有不实，愿意承担一切相关的法律责任。

学位论文作者签名：\_\_\_\_\_

年    月    日