

目录

- 4 基于姿态的手势交互技术.....2
 - 4.1 基于姿态的平面交互手势设计.....2
 - 4.1.1 平面交互基本功能.....3
 - 4.1.2 用户设计手势实验.....4
 - 4.2 平面空间下手势交互实现方法.....9
 - 4.2.1 手部姿态与平面空间坐标映射.....9
 - 4.2.2 基于手部姿态的交互指令实现..... 11
 - 4.3 基于姿态的交互意图识别方法.....13
 - 4.3.1 手部行为震颤与交互意图.....13
 - 4.3.2 交互中手部震颤提取方法..... 15
 - 4.3.3 区分交互意图的实验设计.....18
 - 4.3.4 实验结果与分析.....20
 - 4.4 本章小结.....25

4 基于姿态的手势交互技术

本章旨在探索设计以用户为中心的自然手势交互方式,包括基于姿态的用户设计手势集、构建映射模型和基于手部震颤信息的交互意图识别三个方面。现有研究的手势集无法很好的满足本节基于姿态的平面手势交互需求,而且研究者提出的手势集合可能存在偏见,因此本章首先开展用户参与式的手势设计实验,向用户解释手部姿态的含义,采用大声思考的方式获取用户的想法,最终建立基于姿态的交互手势集合。

针对用户提出的交互手势集合,本章探索了一种坐标映射方法和手势识别方法。根据用户参与式的手势设计实验中,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,提出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。

在手势交互中,难免会出现系统与用户交互意图不匹配的问题,例如用户交互过程中想活动一下手腕,一般来说交互系统会一直将用户行为转化为交互指令,进而出现交互过程中未知的错误。同时,理解用户的交互意图也能提高交互的自然性和流畅性。用户行为包含肌肉的震颤信号,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非常主观行为,不能用简单的线性模型解释,两者都是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进行用户行为的分类。

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;然后针对用户提出的手势集合提出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,证明了使用肌肉震颤判别交互意图的可行性。

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。该研究针对平面手势交互进行设计。第一部分分析了平面交互中的典型功能,确定了平移、选择、抓取、旋转和缩放等交互过程的基本功能。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,包括手势坐标映射与转换为设备指令两步,使各种日常手势都可以直观地用来控制平面交互过程,实现了一套符合用户习惯和偏好且支持多种典型操作功能的平面手势交互集。

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【1】，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。结合【西北工业大学 王阳阳、王刚】等人的研究，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【1】。

这五大基本手势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【1】

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【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,因此设计平移手势尤为重要。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。

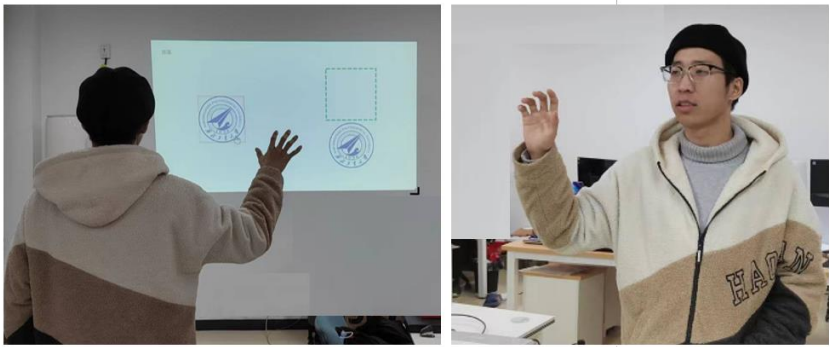
尽管先前研究提出了各种基于手掌和手指姿态的手势集【2】，但经过考量,本研究认为仅依靠手部姿态难以有效实现平面交互所需的这些功能，尤其平移功能无法得到很好的映射，例如用户执行平移向左的功能时，只需要将手从左边移动到右边，在这个移动的过程中，用户的手部姿态不会发生变化。

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,本研究将采用以用户为中心的构建手势集实验。

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,目的设计一组以用户为中心且基于姿态的手势集合。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。

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。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。实验过程中,要求用户采用大声思考法【1】,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。



【贴张图，交互实验的图】

参与者与实验环境:

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~25 岁【(均值=23, SD=1.5)】,其中 6 人惯用手为右手,3 名参与者有空中手势交互的经验,4 名参与者均没有空中手势交互的经验。实验在所用设备为英特尔 i7-6700、英伟达 GTX960 的电脑,投影仪为惠普型号,投影尺寸为【1】。

实验设计:

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

【】，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。在实验中，用户站在大屏前，根据提示进行手势设计，每个功能需要完整的做一遍以确定选择的手势动作，实验全程记录用户的设计手势与思考内容。在表【】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【】 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	

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。

在本研究中，引入手势集合的一致性系数【】对手势集合进行分析，手势的一致性系数定义如公式【】。

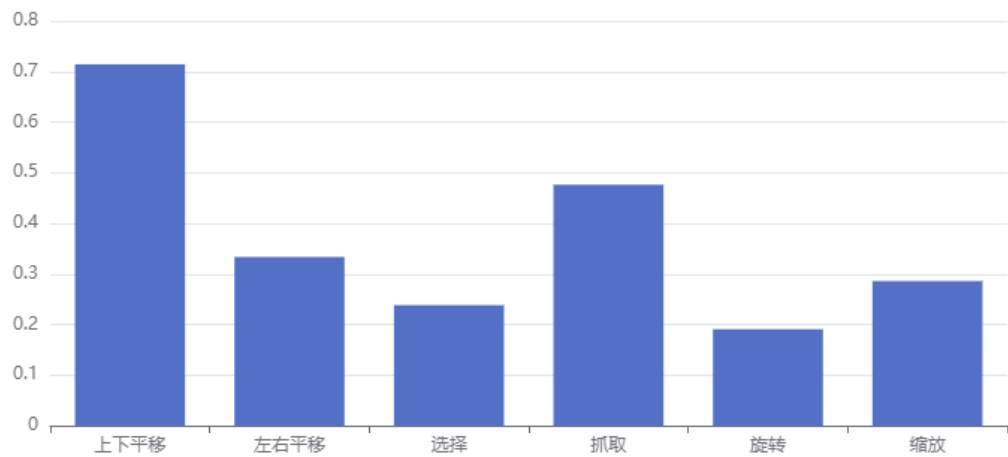
$$\mathcal{AR}(r) = \frac{|P|}{|P| - 1} \sum_{P_i \subseteq P} \left(\frac{|P_i|}{|P|} \right)^2 - \frac{1}{|P| - 1}$$

在公式【】中，P 是参与者的数量，Pi 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，该手势从用户体验角度来看更自然、更容易被接受，该手势获得了更高层次的用户共识,更适合作为设计输出。

实验结果：

实验总共收集了 9x10=90 个手势，根据收集到的手势特征，对具有相同特点的手势进行合并处理，将其合并为同一类手势，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。

根据收集到的手势集合，根据公式【】，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【】所示。



图【】 用户参与设计手势的一致率分数

在图【】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。

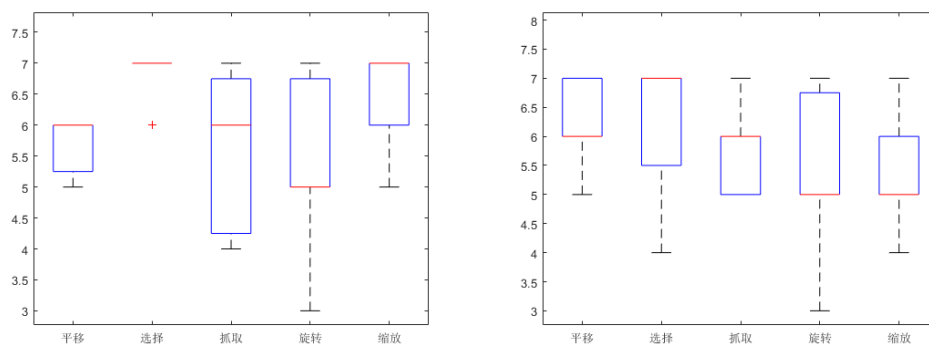
在选择这一手势上，用户几乎都没有过多思考就给出了手势，有三名用户使用食指点击进行选择，有一名用户说“这个手势就像点鼠标一样，点一下就选择了”，有两名用户使用食指和拇指捏住的手势，有一名用户说“选择可能会很频繁的触发，因此我选择食指与拇指捏住这个手势，因为这个很容易执行”，有两名用户选择握拳作为选择手势。

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，绘制出盒图如图【】所示。



a) 【易用性】 b) 【直观性】

【】 用户手势评分箱线图







在图【】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，都认为自己针对选择提出的手势容易执行。针对平移手势则用户【】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，所以打了5分。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。

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集存在冲突时,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自主性之间取得最优解，手势集合见表【】。

表【】 平移、点选、抓取、旋转、缩放

交互功能		功能描述	手势
平移	上	向上移动	

	下	向下移动	
	左	向左移动	
	右	向右移动	
点选		选择对象	
抓取		抓起对象	
旋转	逆时针	使对象逆时针旋转	
	顺时针	使对象逆时针旋转	
缩放	放大	使对象放大	
	缩小	使对象缩小	

4.2 平面空间下手势交互实现方法

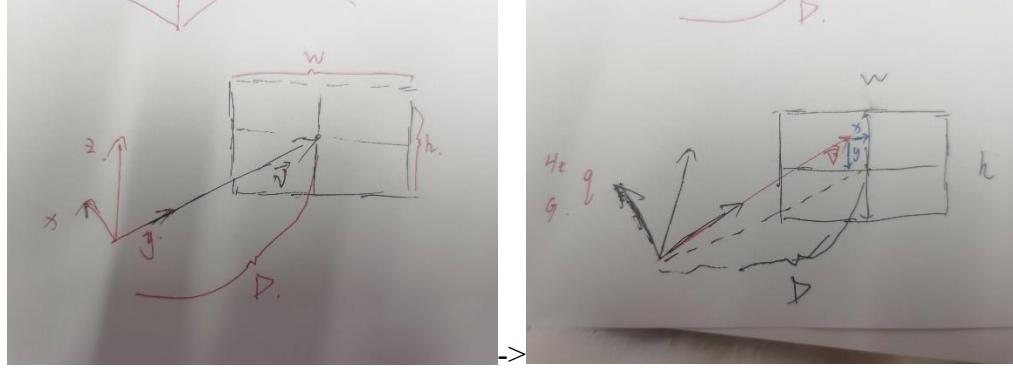
本节主要内容为探究以用户为中心，基于姿态的手势集合实现方法，主要分为坐标映射与手势指令实现两部分。

全手行为跟踪方法是大屏手势交互的核心部分，然而在大屏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。然后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。



【b 图不规范的，重新画】

a) 旋转前姿态与向量映射图 b) 旋转后姿态与向量映射图

图【】 坐标系转换模型

利用手部姿态，将以用户手所在向量为 ${}^t v = [0, D, 0]$ ，当用户在单位时间内执行俯仰动作且转动的角度为 θ ，该过程用四元数 ${}^L q$ ，执行偏航动作且转动角度为 φ 时，该过程用四元数 ${}^{H_{t+1}}_L q$ 表示，该转动过程的姿态四元数满足公式【】，用户手所在向量 ${}^t v$ 的转动结果如公式【】。

$${}^{H_{t+1}}_{H_t} q = {}^{H_{t+1}}_L q \otimes {}^L q \quad \text{公式【】}$$

$${}^{t+1} v = R({}^{H_{t+1}}_{H_t} q) {}^t v \quad \text{公式【】}$$

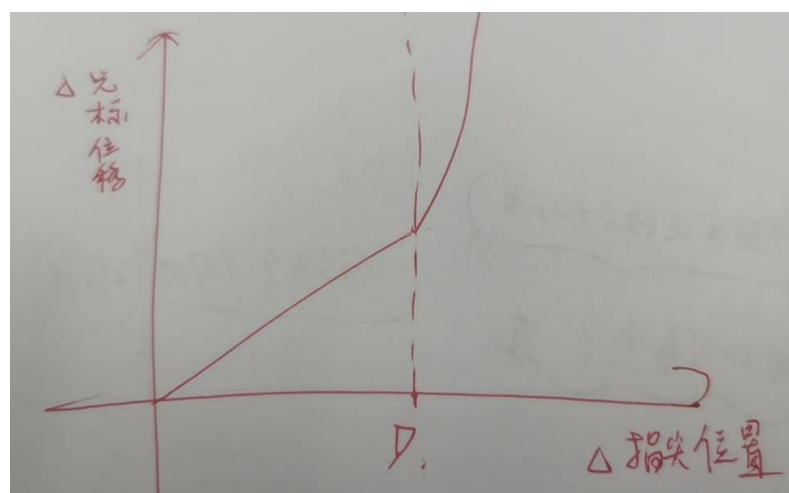
将公式【】向量的 ${}^{t+1} v$ 的 x 分量和 z 分量的值映射到平面空间中竖直方向的移动和水平方向的移动。这种映射方式受到 4.1 节中用户的启发，用户选择手掌俯仰转动作为控制光标上下移动的手势时，提出根据指尖的位置进行控制。鉴于不同的手势需要手指运动实现，因此本研究将其抽象为将中指掌骨视为移动的向量，掌骨末端相较于手腕的空间移动，映射为光标的空间移动。

$${}^{H_t}_{H_{t-1}} q = {}^{H_t}_G q \otimes ({}^{H_{t-1}}_G q_{mag})^{-1}$$

在公式【】中，【】。该过程为线性映射过程，上述过程将手部姿态的变化角度映射为了 2D 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,如 proportion mapping。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。

基于 4.1 【】 用户提出的担忧,本文改进了这一映射方案。有研究者【】提出了一种 Go-Go 交互技术,使用交互式增长了用户的非线性映射。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。



图【】 基于分段函数的非线性映射图

在图【】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。这表明在映射时是一段分段函数,有研究者【】指出人手腕关节的俯仰角度是 110 度的范围,偏航却只有 30-50 度的范围,这为本研究设计分段函数的临界点与函数提供了参考。同时有研究者指出【】,人的一个行为所需时间约为 1 秒钟,因此建立以人为中心的空间映射模型。综合考虑,本研究建立映射分段函数,如公式【】所示。

$$\Delta C = \begin{cases} a \times \Delta f, & \Delta f < D \\ a \times \Delta f + (\Delta f - D)^2, & \text{other} \end{cases}$$

在公式【】中, ΔC 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, Δf 表示指尖相对于掌心移动距离变化量, 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【】。

表【】 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【】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，但这样更容易误识别，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,我们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。然后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,我们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【】所示。

在图【】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。

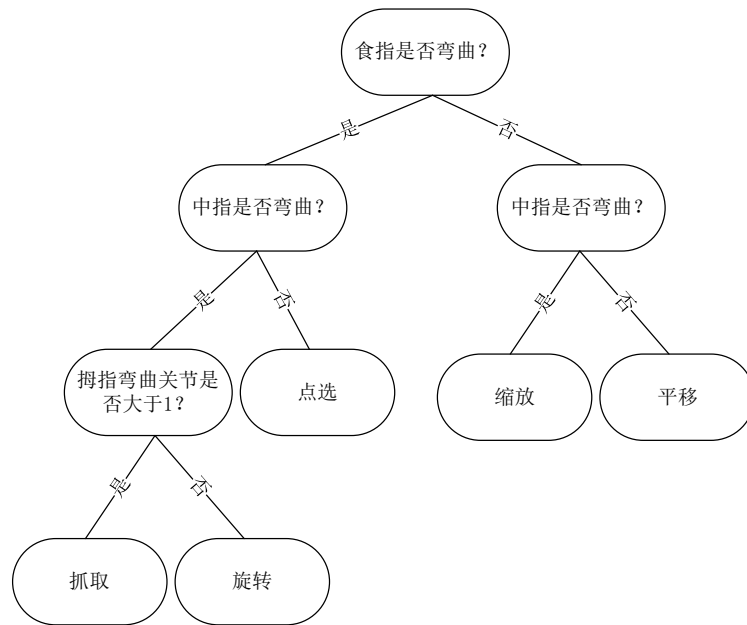


图 【】 手势识别的决策树结构

4.3 基于姿态的交互意图识别方法

本小结主要探究含有交互意图的手部震颤特征。通过震颤的特性挖掘震颤信号的特征，判断与交互意图之间的关系。本章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，而交互意图则并不能用简单的线性随机模型进行描述，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。然后基于震颤振幅、频率等时间/频域特征以及符合线性随机过程的特点，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，准确率高达 96%。但仅通过震颤特征很难区分不同难度任务中动作性震颤的差异。因此该研究验证了人体震颤与交互意图之间存在关联。

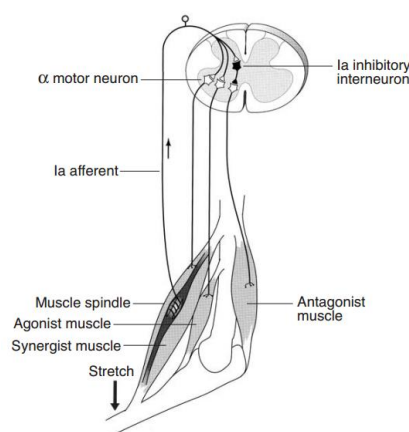
本章旨在通过健康人群的震颤信号特征，分析用户的交互意图，以实现更自然的手势交互。

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。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【4.3.3 小节】的实验设计中，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，判断与交互意图之间的关系。

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。

自 1993 年以来,有两种主要的震颤分类在使用。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,因此采用另一种分类方式。而根据 Bain (1993)和 Deuschl 等人(1998)的报告,第一种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。本论文基于该分类方式探究震颤特征与交互意图之前的联系。



图【】 震颤原理图

研究者 Elble 等人【】(1986)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。研究者 Deuschl 等人【】表明静息性震颤频率通常在 4hz-6hz, 而姿势性震颤频率通常在 4hz-12hz 之间。研究者 Elble 和 Koller【】等人指出正常的生理震颤通常集中在 8-12 赫兹的范围,但不同人也可能存在较大差异。同时,研究者【】指出震颤是一种线性随机过程。

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、

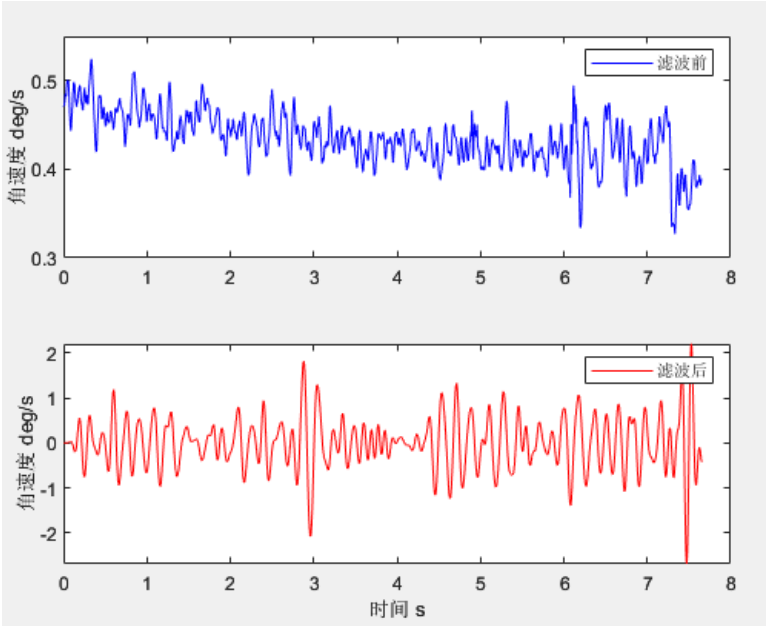
中、后,用户的意图也会不断发生变化。这同样是个动态调整的过程,难以用静态的线性关系解释。不同用户由于其个体差异,在相似环境和目标下,会产生不同的交互策略,这种多样性也决定了非线性的特点。用户意图表达和意图理解中的语言、表情、手势等都是复杂的交互要素,互相影响,也呈现出强非线性。

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

在本小节中,根据震颤信号的振幅、频率和符合线性随机过程的特性,描述了震颤信号提取方法,震颤的偏度和峰度特征和具有时间反演对称性,同时构建了 14 个特征以进行实验设计与分析。

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振幅(平均振幅)。如果信号平均值为零且无整体漂移趋势(如震颤加速度记录),那么振幅即为信号的均方根,也基本等于标准差。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 赫兹,并且一般不高于 12hz。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【】所示。



图【】

在图【】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(4hz-6hz)与动作性震颤信号(6hz-12hz)。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。

有研究者【】指出震颤信号是一种线性随机过程(Timmer et al 2000)，病理性震颤更为非线性。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。

如果生理性震颤具有高斯随机性且基本过程为线性,那么检测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)反映分布的对称性,第四阶矩(峰度)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【】给出了偏度的计算过程。

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

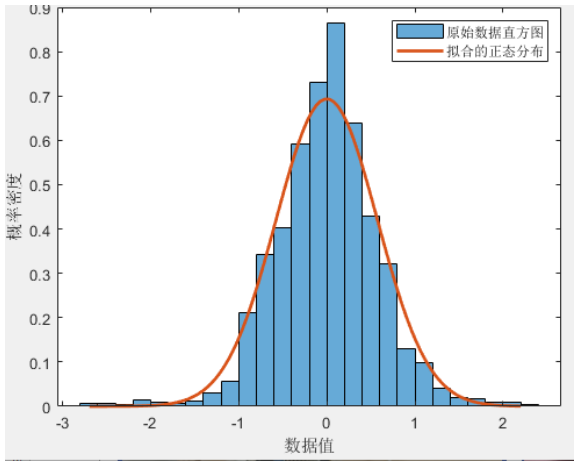
公式【】为偏度计算公式，其中 n 表示数据总数， X_i 是第 i 个观察值， s 表示观察值的标准差， \bar{X} 表示观察值的均值。偏度是描述概率分布形状对称性的一个统计量。它反映分布曲线相对于中心的左右尾部延伸的情况。偏度为正表示右侧尾部更长,负偏度则表示左侧尾部更长。偏度绝对值越大表示分布曲线越不对称。零偏度对应的是对称分布,如正态分布。判断分布的偏斜方向和程度,偏度是最有效的度量标准,可判断实际分布与理论对称分布的差异。

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。

$$m_4 = \frac{1}{n} \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{s} \right)^4 - 3$$

公式【】为峰度计算公式，其中 n 表示数据总数， X_i 是第 i 个观察值， s 表示观察值的标准差， \bar{X} 表示观察值的均值。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【】所示。



图【】健康人群震颤信号分布直方图

有研究者指出【Edwards and Beuter 2000】，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【】。

$$\max\{|R_j|, \text{time lag } j\}, \text{ where } R_j = \frac{\sum_i (x_i - x_{i-j})^3}{(\sum_i x_i^2)^{3/2}}.$$

在公式【】中,lagj 为时间滞后参数, x 为分布中的变量。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。

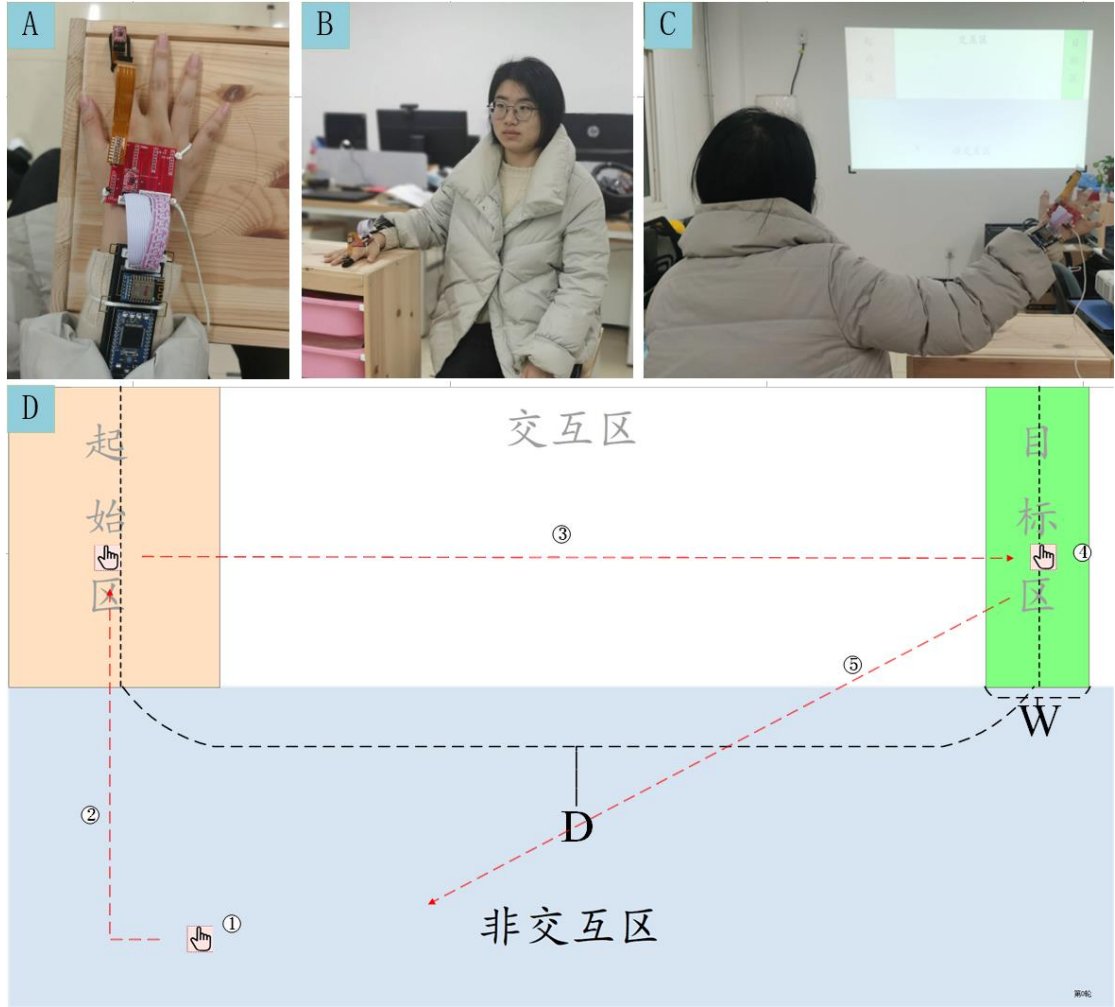
如上所述,震颤信号的线性随机过程所具有的特征可以有效判断信号的线性与非线性特性,。频谱分析的常用方法还包括功率谱密度分析和主频率分析。功率谱密度反映信号功率在频域上的分布情况。通过离散傅里叶变换可以计算得到静息态和运动态震颤信号的功率谱密度值,作为判断震颤强度的指标。主频率分析即识别信号中占主导地位的频率成分。通常震颤信号会包含一个主要的频率成分,这可以反映震颤的类型和状态。短时傅里叶变换的时频分析可以有效提取该主频率。因此,频谱分析中功率谱密度和主频率等参数的计算,可以从频域角度更全面地分析判断震颤的特征。综合研究者【】,本文建立了包括时域特征,见表【】。

表 【】 构建特征集合

特征类别	特征
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)
	线性对称性(R)
频域特征	4-6Hz 功率谱密度(LPSD)
	6-12Hz 功率谱密度(HPSD)
	主频率 (FT)
线性随机过程特征	偏度(m3)
	峰度(m4)

4.3.3 区分交互意图的实验设计

本小节旨在探究健康人群在各种交互任务中交互意图与用户行为震颤的联系。因此本实验设计不同难度的交互任务并收集 RT、IAT、PAT 信号，分析其震颤特征。本实验交互实验中采集用户的震颤信号，分别在非交互区用户将手放在桌子上静止、从起始区移动到目标区、在目标区保持静止的震颤信号，交互任务见图【】。本实验根据 Fitts 定律模型^{【…】}构建交互任务，根据难度指数（ID）的 Shannon^{【引用】}公式【】给定了固定距离 D 下三种不同难度系数。为了去除无关变量的干扰，本研究给予用户明确的交互任务，使用手部姿态尽快尽准确的控制光标从起始区移动到目标区。在交互过程中采集用户 RT、IAT、PAT 信号旨在探究用户无交互意图下震颤信号与含有交互意图震颤特征的区别。



A)交互设备及佩戴分布 B)用户实验环境 C)交互过程 D)实验任务设计图

图【】实验环境与设计图

根据【4.3.2 小结】震颤信号受到许多因素的影响，例如吸烟、咖啡等，本研究旨在探究健康人群的震颤特征，因此也需要将其作为考虑，在征集用户信息阶段，针对参与者征集其年龄；性别；受教育水平；惯用手；锻炼、咖啡、吸烟、饮酒的频率及最近一次的时间；手部脑部疾病史；近期是否服药，治疗何种病和药品名称；收集这些全面的人口统计信息可以对数据进行更细致的分析。数据有助于理解不同的因素如何影响健康个体的意图震颤，理解健康用户的震颤特征，还有助于控制分析中潜在的混杂变量。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。

$$ID = \log_2 \left(\frac{D}{W} + 1 \right),$$

公式【】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。

参与者与实验环境：

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($\text{mean}=20.55, \text{SD}=1.013$)，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。

实验使用惠普【】投影仪投影到【xx 寸】进行交互，数据收集装置见图【-a】，通过附着在手上的两个惯性传感器进行数据采集和系统控制，手背传感器负责控制光标移动、食指指尖传感器负责采集用户的震颤信号，采集三轴加速度与三轴角速度。

实验设计：

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 个不同的目标区宽度的任务，每个任务有 3 个 block，用户每个 block 中需要执行 10 次，共 $10 \times 3 \times 3 \times 9 = 810$ 条数据，为了消除交互任务顺序带来的影响，本实验采用拉丁方的方法进行实验，实验过程中需要用户坐在椅子上，将手放在实验桌上，以便于实验时采集静止数据。

实验的独立变量为目标区宽度 W：30pt、90pt、180pt。（难度指数 ID：5.67、4.14、3.22）。实验过程中的反馈分为视觉反馈和听觉反馈。视觉反馈为交互区域高亮，提示用户需要将光标移动/停留在高亮区域内；听力反馈为交互过程中的叮咚声，叮咚声提示用户要开始移动光标。

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；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，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。在实验开始前，明确告知用户所有的声音提示都是开始运动的提示，所有的高亮提示都表示需要高亮区域停留。

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。

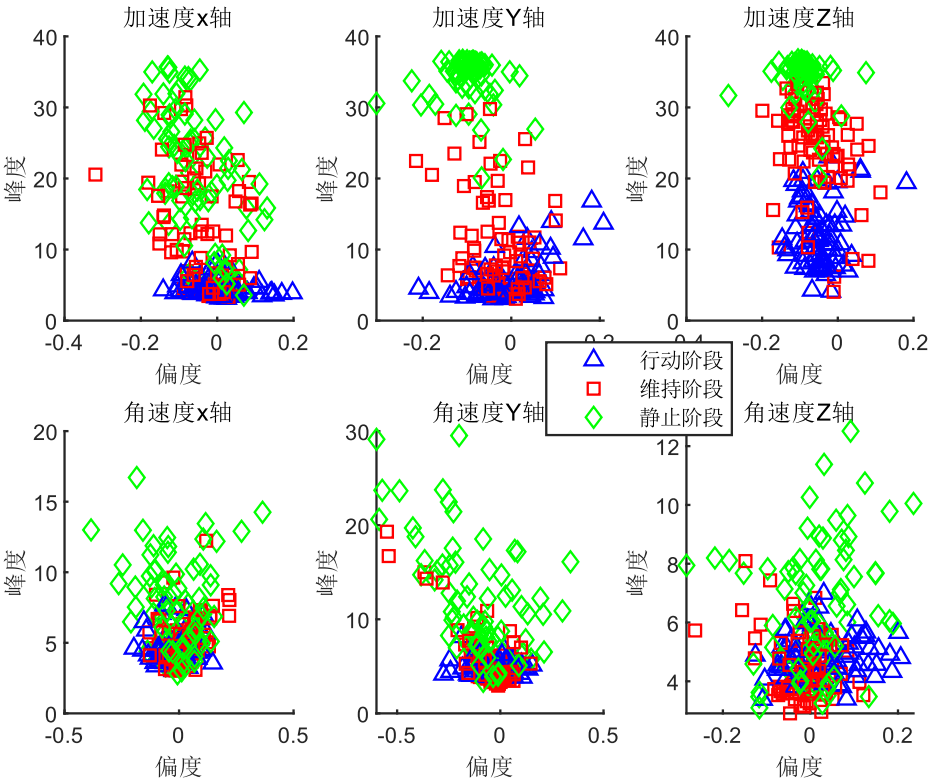
4.3.4 实验结果与分析

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 m_3 、 m_4 特征、 R 特征以及 LPSD、HPSD 特征进行分析，并提取【4.3.1 小节中】构造的特征，使用支持向量机（SVM）、 k 最近邻（KNN）、决策树（Decision Tree）、朴素贝叶斯（Naive Bayes）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。

数据预处理：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(4hz-6hz)与动作性震颤信号($\text{Sig} < 12\text{hz}$)。根据标记和幅值检测，删除一些无效的数据，经过筛选后，有效数据数目为 785 条。

提取峰度(m_3)、偏度(m_4)特征：通过【4.3.1 小节】所述公式，提取六轴信息的 m_3 、 m_4 特征，以实验中 block 为单位，共计得到 RT、IAT、PAT 各 27 条数据，以偏度为横轴，峰度为纵轴绘制分布图，见图【】。

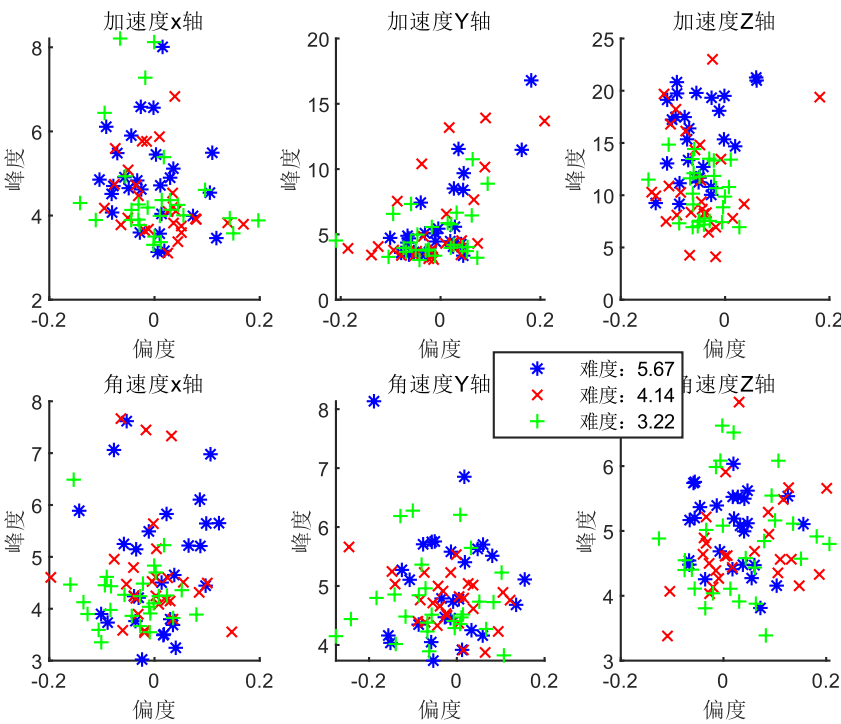


图【】 不同震颤类别的峰度偏度分布图

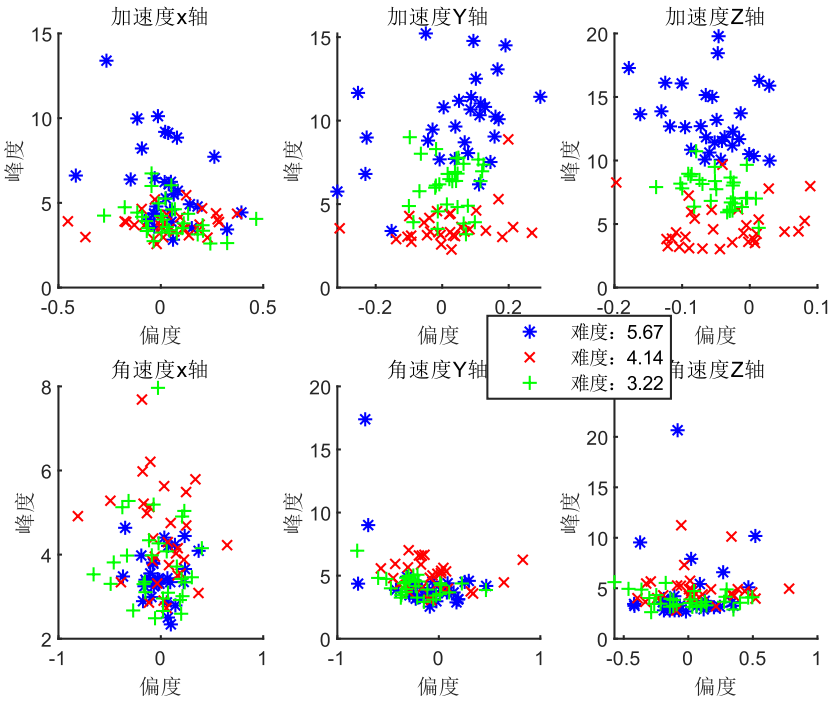
在图【】中，三种震颤信号的峰度在加速度 Y 轴与 Z 轴表现出差异显著，通过 Wilcoxon 秩和检验中 RT 与 PAT 具有显著差异($Z = -9.12, p < 0.01$)，RT 与 IAT 具有显著差异($Z = -10.9, p < 0.01$)，在六轴数据中，RT 的峰度均值高于 PAT 和 IAT，PAT 和 IAT 在加速度 Z 轴差异显著($Z = -10.2, p < 0.01$)。

在上述分布图中发现，通过加速度 Z 轴与 Y 轴的峰度偏度分布图能够区分 RT、IAT、PAT，为了进一步探究能够通过震颤信号的峰度偏度分布区分不同任

务难度下的 IAT，以 block 为单位，计算六轴数据的 m3、m4 特征，绘制分布图【】。



a) 所有用户峰度偏度分布图



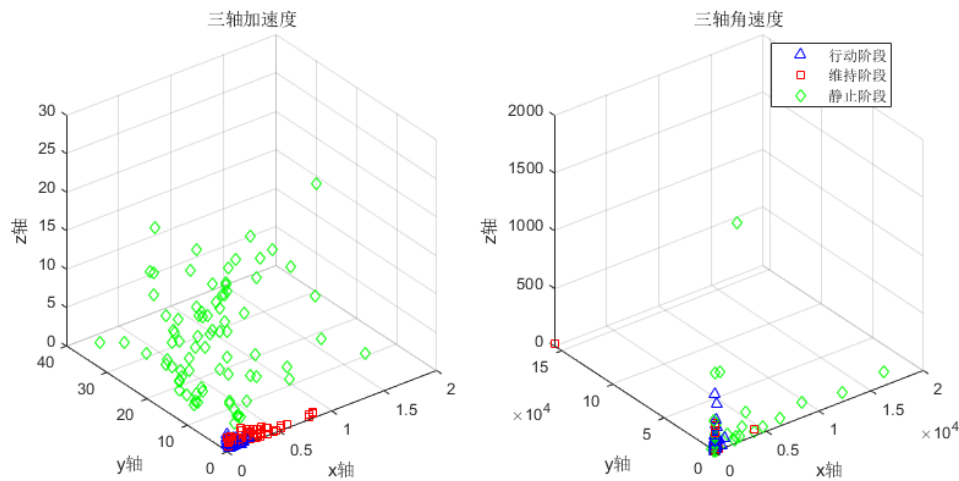
b) 用户 3 的峰度偏度分布图

图【】 不同任务难度下运动性震颤的峰度偏度分布图

通过图【a】发现六轴数据中并未出现分层现象，通过六轴震颤信号的峰度偏度特征无法对所有人的交互意图进行分别。为了探究是否由于个体的差异导致震颤信号的不可区分，本实验以每个人的交互任务为单位，计算六轴数据的m3、m4特征，绘制用户的峰度偏度分布图，图【b】是用户3的分布示例图，在图中得出加速度Y和Z轴的特征出现分层现象，通过单因素方差分析可知在加速度Z轴($F_{2,27} = 135.4, p < 0.01$)和加速度Y轴($F_{2,27} = 80.2, p < 0.01$)上均具有显著差异。

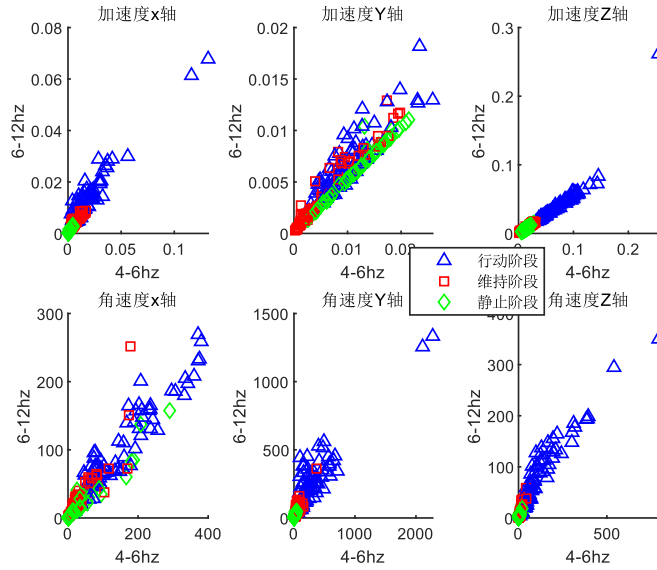
峰度表示为信号分布的尖峭度，通过图【】可以得出，RT在加速度Z轴与Y轴上峰值最高，而PAT的峰值次之，IAT信号峰值最低。而利用峰度偏度对IAT信号进行区分时发现，分布图中并未出现分层，但针对用户3三种难度下的IAT出现了显著差异。

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，但无法区分不同难度下的IAT。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。



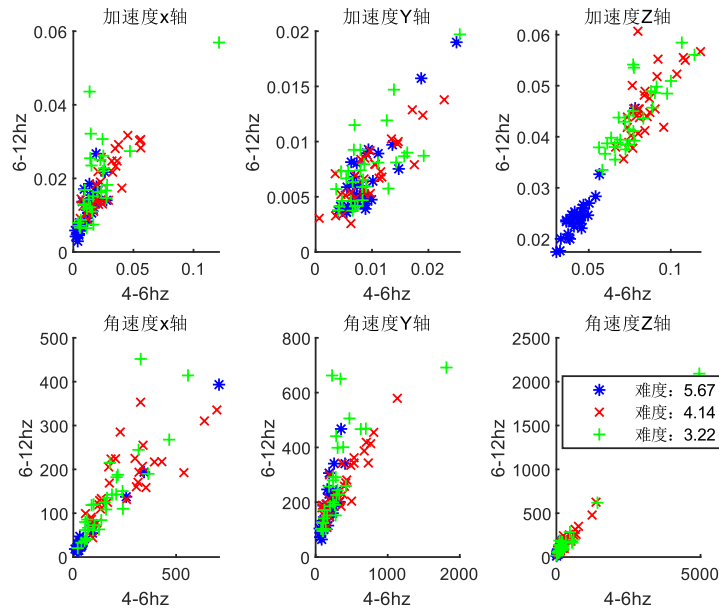
图【】不同震颤类别的三轴分布图

在图【】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【】。



图【】 不同震颤类别的 LPSD-HPSD 分布图

在图【】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。为了进一步探究 LPSD-HPSD 分布能否分辨出不同任务难度下的 IAT，本文绘制了不同任务难度下运动性震颤的 LPSD-HPSD 分布图，见图【】。



图【】 不同任务难度下运动性震颤的 LPSD-HPSD 分布图

通过图【】可以看出，通过加速度 z 轴中 IAT 的特征可以区分难度 5.67 任务与其他难度，均具有较低的功率谱密度，这表明通过加速度 Z 轴的 HPSPD 和 LPSD 能够区分最高难度与其他难度下的 IAT。

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，但无法对不同难度任务的 IAT 进行很好的区分，本实验结合【4.1.1】

提出的特征集合，分别从六轴中提取特征，使用 SVM、KNN、Decision Tree 和 Naive Bayes 进行分类。表【】给出了 RT、IAT、PAT 分类的准确率、F1 分数和召回率，表【】给出了三种难度下 IAT 分类的准确率、F1 分数和召回率。

表【】 区分三种难度下 IAT 结果

分类模型	Accuracy	F1 Score	Recall
SVM	0.469	0.548	0.525
KNN	0.456	0.547	0.542
Decision Tree	0.543	0.616	0.627
Naive Bayes	0.531	0.684	0.864

表【】 区分 RT、IAT、PAT 结果

分类模型	Accuracy	F1 Score	Recall
SVM	0.573	0.641	0.487
KNN	0.900	0.949	0.943
Decision Tree	0.962	0.996	1
Naive Bayes	0.943	0.987	0.975

通过分析不同分类器的结果发现 Decision Tree 分类器不论对 RT、IAT、PAT 分类还是对三种难度下 IAT 分类都取得了最高的准确率。在通过震颤特征对 RT、IAT、PAT 分类时，准确率最高为 96%（Decision Tree）。而针对三种难度下 ICA 分类准确率最高为 54%（Decision Tree）。

综上所述，通过震颤信号的特征，区分 IAT 与不含有交互意图的震颤信号具有较高的准确率（96%），而对于不同难度下 IAT 的分类则最高只有 54%的准确率。因此通过震颤能够区分有交互意图行为和无交互意图行为的运动性震颤，而对于不同交互意图的运动性震颤仅仅只有 54%的准确率。

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,通过用户参与式设计实验,在向用户解释手部姿态含义的基础上,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设计者主观偏见的问题。其次,针对用户设计的手势集合,提出了一种坐标映射方法和手势识别方法。坐标映射方法解决了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;手势识别方法则通过决策树对手指关节特征进行分类。最后,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，Decision Tree

分类器达到了 96%的准确率，区分不同任务难度下的 IAT 时，也有 54%的准确率，这证明了使用肌肉震颤识别交互意图的可行性。

总的来说,本章内容从手势交互的需求分析出发,通过用户实验获得设计基础,在此基础上探索相关技术方法,并从理论上确立这些方法的有效性。