

基于最优控制理论的触摸运动模型 及交互优化技术

(申请清华大学工学博士学位论文)

培养单位：计算机科学与技术系

学 科：计算机科学与技术

研 究 生：古 肇 正

指 导 教 师：史 元 春 教 授

副指导教师：喻 纯 副教授

二〇二二年三月

Modeling, Sensing and Understanding Pervasive Touch Input

Thesis Submitted to
Tsinghua University
in partial fulfillment of the requirement
for the degree of
Doctor of Philosophy
in
Computer Science and Technology

by

Gu Yizheng

Thesis Supervisor: Professor Shi Yuanchun

Associate Supervisor: Associate Professor Yu Chun

March, 2022

学位论文指导小组、公开评阅人和答辩委员会名单

指导小组名单

李 XX	教授	清华大学
王 XX	副教授	清华大学
张 XX	助理教授	清华大学

公开评阅人名单

刘 XX	教授	清华大学
陈 XX	副教授	XXXX 大学
杨 XX	研究员	中国 XXXX 科学院 XXXXXXX 研究所

答辩委员会名单

主席	赵 XX	教授	清华大学
委员	刘 XX	教授	清华大学
	杨 XX	研究员	中国 XXXX 科学院 XXXXXXX 研究所
	黄 XX	教授	XXXX 大学
	周 XX	副教授	XXXX 大学
秘书	吴 XX	助理研究员	清华大学

关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：

清华大学拥有在著作权法规定范围内学位论文的使用权，其中包括：（1）已获学位的研究生必须按学校规定提交学位论文，学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文；（2）为教学和科研目的，学校可以将公开的学位论文作为资料在图书馆、资料室等场所供校内师生阅读，或在校园网上供校内师生浏览部分内容；（3）按照上级教育主管部门督导、抽查等要求，报送相应的学位论文。

本人保证遵守上述规定。

作者签名：_____

导师签名：_____

日 期：_____

日 期：_____

摘要

触摸交互是最重要的人机交互方式之一，是人控制手指触摸交互表面，通过点击、长按、滑动等手势向计算机输入信息的方式。触摸交互有三大优化目标：（1）普适性：人渴望随时随地与计算机交互，触摸交互不应局限于有源表面，而应拓展至无源表面（非数字化表面，如桌面、墙面）；（2）响应性：人对触摸交互响应性的感官需求极高，最低能察觉到 24 毫秒的延迟^[1]，触摸交互的响应性优化应以用户无法察觉延迟为目标；（3）意图性：自然触摸交互中人的有意触摸和无意触碰混杂，防误触技术应从用户意图层面定义误触，彻底解决误触问题。尽管触摸屏已经问世数十年，它在普适性、响应性和意图性上均未达标；不少研究工作针对上述问题改进触摸交互技术，但始终缺少基本原理作为指导。在此背景下，本文提出触摸运动模型，揭示触摸前后极短时间内手指的运动规律，为基于运动传感的触摸交互技术提供了计算理论基础。基于模型，本文针对多种触摸交互任务（目标选择、文本输入），优化触摸感知和防误触技术，提升交互效率和用户体验。具体贡献如下：

（1）提出了基于最优控制理论的触摸运动模型，包含数学表达（描述触摸运动的数学方程）和计算方法（利用位移、速度、加速度传感信号拟合触摸运动方程的计算过程）。本文详细描述了模型及其推导过程，通过实验验证了模型的有效性，最后讨论了模型在普适性、响应性和意图性三方面对触摸交互技术的指导意义。

（2）提出了基于惯性传感指环的低延迟触摸感知技术：指环触摸感知技术让人能在无源表面上触摸点击输入，先前工作通过阈值方法感知触摸^[2]，延迟高达 200 毫秒，准确率仅为 85%。本文基于触摸运动模型，提出了低延迟的指环触摸感知技术，延迟低于 10 毫秒，准确率超过 99%。由于人无法察觉到低于 10 毫秒的触摸交互延迟，该技术达到了最佳的响应性。

（3）提出了基于惯性传感指环的无源表面文本输入技术：通过拓展触摸运动模型，惯性传感指环不仅能感知无源表面上的触摸点击，还能识别包括长按和滑动在内的触摸交互手势。综合触摸手势识别能力，本文提出无源表面文本输入技术，该技术下人平均每分钟可输入 20.6 个英文单词，效率超出先前可比技术 40.7%。

（4）提出了面向平板电脑文本输入的防误触技术：平板电脑上的文本输入是最复杂的连续触摸输入任务，针对平板十指文本输入中误触频发的问题，本文提出基于触摸运动模型的防误触技术。本文通过文本输入任务评估防误触技术的性能，防误触技术的准确率达到 99%，且允许用户在打字时将非交互手指休息在触

摘 要

摸屏上（而不会引发误触），将打字速率提升 20%。

关键词：触摸交互；感知技术；意图识别

Abstract

An abstract of a dissertation is a summary and extraction of research work and contributions. Included in an abstract should be description of research topic and research objective, brief introduction to methodology and research process, and summary of conclusion and contributions of the research. An abstract should be characterized by independence and clarity and carry identical information with the dissertation. It should be such that the general idea and major contributions of the dissertation are conveyed without reading the dissertation.

An abstract should be concise and to the point. It is a misunderstanding to make an abstract an outline of the dissertation and words “the first chapter”, “the second chapter” and the like should be avoided in the abstract.

Keywords are terms used in a dissertation for indexing, reflecting core information of the dissertation. An abstract may contain a maximum of 5 keywords, with semi-colons used in between to separate one another.

Keywords: keyword 1; keyword 2; keyword 3; keyword 4; keyword 5

目 录

摘要.....	I
Abstract.....	III
目录.....	IV
插图和附表清单.....	VIII
符号和缩略语说明.....	XI
第1章 引言	1
1.1 选题背景及意义	1
1.1.1 触摸交互的重要性.....	1
1.1.2 触摸交互的改进空间.....	2
1.1.3 为什么要建立触摸运动模型?	3
1.2 研究现状	4
1.2.1 触摸交互技术.....	4
1.2.2 触摸交互模型.....	5
1.3 研究内容	6
1.3.1 触摸运动模型的数学表达.....	7
1.3.2 触摸运动模型的计算方法.....	7
1.3.3 基于触摸运动模型的交互优化技术.....	8
1.4 主要研究成果	9
1.5 论文组织结构	11
第2章 基于最优控制理论的触摸运动模型	12
2.1 触摸运动模型的数学表达	12
2.1.1 单位触摸运动方程.....	14
2.1.2 触摸运动方程的推导过程.....	15
2.1.3 更复杂的触摸运动方程.....	16
2.2 触摸运动模型的计算方法	17
2.2.1 卡尔曼滤波.....	17
2.2.2 最小二乘拟合	19

2.3 触摸运动模型的实验评估	20
2.3.1 数据采集实验	20
2.3.2 无约束运动模型的合理性	24
2.3.3 触摸运动模型的拟合精度	24
2.4 触摸运动模型的指导意义	26
2.4.1 触摸检测	27
2.4.2 触摸手势识别	28
2.4.3 触摸意图推理	29
2.5 本章小结	31
第 3 章 指环上的高准确低延迟触摸检测技术	32
3.1 引言	32
3.2 相关工作	34
3.2.1 基于视觉的无源表面触摸交互	34
3.2.2 基于震动的无源表面触摸交互	34
3.2.3 触摸姿态	35
3.3 指环触摸交互的自然性研究	35
3.3.1 实验一 A: 触摸姿态主观喜好程度调研	35
3.3.2 实验一 B: 指环佩戴位置主观喜好程度调研	38
3.4 基于机器学习的触摸检测技术	39
3.4.1 实验二: 收集触摸数据	39
3.4.2 基于机器学习的触摸姿态分类	42
3.4.3 基于机器学习的触摸检测技术	42
3.5 基于触摸运动模型的触摸检测技术	49
3.5.1 基于触摸运动模型的触摸检测技术	49
3.5.2 实验三: 评测本节触摸检测技术	52
3.6 本章小结	54
第 4 章 指环上的触摸手势识别与打字技术	55
4.1 引言	55
4.2 相关工作	57
4.2.1 可穿戴文本输入	57
4.2.2 智能指环交互	59
4.2.3 贝叶斯解码器	60

4.3 智能打字指环的交互设计	61
4.3.1 键盘布局	61
4.3.2 交互流程	62
4.3.3 设计准则	62
4.4 指环上的触摸手势识别	63
4.4.1 实验一：收集触摸手势数据	63
4.4.2 触摸手势识别技术	65
4.5 智能打字指环的解码器设计	72
4.5.1 实验二：收集打字数据	72
4.5.2 智能打字指环单词解码器	76
4.6 智能打字指环的评测	79
4.6.1 实验三：智能打字指环的评测实验	79
4.6.2 智能打字指环的评测结果	80
4.6.3 与手机打字速度进行比较	85
4.6.4 对评测结果的讨论	86
4.7 本章小结	89
第 5 章 平板电脑上的触摸意图推理键盘	90
5.1 引言	90
5.2 相关工作	92
5.2.1 触摸交互中的防误触技术	92
5.2.2 触摸交互防误触技术的作用	94
5.3 初版防误触键盘	95
5.3.1 实验一：收集无反馈键盘上的打字数据	95
5.3.2 基于打字行为设计初版防误触键盘	98
5.4 终版防误触键盘	103
5.4.1 实验二：收集初版防误触键盘上的打字数据	103
5.4.2 基于打字行为设计终版防误触键盘	104
5.4.3 关于防误触键盘的讨论	106
5.5 防误触键盘的评测	110
5.5.1 实验三：防误触键盘的评测实验	110
5.5.2 防误触键盘的评测结果	111
5.6 本章小结	116

目 录

第 6 章 结论和展望	117
6.1 结论	117
6.2 展望	117
参考文献	118
附录 A 补充内容	131
致 谢	132
声 明	133
个人简历、在学期间完成的相关学术成果	134
指导教师学术评语	136
答辩委员会决议书	137

插图和附表清单

图 1.1 无源表面触摸交互的构想	2
图 1.2 无源表面触摸交互的相关论文	3
图 2.1 无约束运动模型	13
图 2.2 触摸运动模型	14
图 2.3 单位触摸运动方程	15
图 2.4 常用的运动传感器	17
图 2.5 触摸运动模型评估实验设置	21
图 2.6 触摸踏空实验	22
图 2.7 触摸运动模型的拟合实例	25
图 2.8 抛物线对触摸运动的拟合实例	26
图 2.9 模型为触摸检测提供判断依据	27
图 2.10 模型为抬起检测提供辅助的判断依据	28
图 2.11 触摸速度与力度之间的关系	30
图 3.1 MR 场景下触摸交互无处不在的设想	33
图 3.2 触摸姿态的三维设计空间	36
图 3.3 触摸姿态主观评分热力图	37
图 3.4 人们最喜好的触摸姿态前十名	38
图 3.5 指环佩戴位置主观评分热力图	39
图 3.6 实验设置	41
图 3.7 加速度信号图示	43
图 3.8 角速度和重力信号图示	44
图 3.9 触摸检测分类器	45
图 3.10 触摸检测准确率与延迟之间的关系	46
图 3.11 分类器对各触摸姿态的检测准确率	47
图 3.12 本触摸检测技术与 baseline 的对比	49
图 3.13 实验任务图示	53
图 4.1 智能打字指环的交互流程	55
图 4.2 智能打字指环的键盘布局	61
图 4.3 实验一的实验设置	64
图 4.4 触摸与抬起速度之间的关系	67

图 4.5 触摸交互信号图示	68
图 4.6 点击、左滑、右滑三分类模型	70
图 4.7 实验设置	71
图 4.8 降采样情况下触摸手势识别的准确率	72
图 4.9 实验二的实验设置	73
图 4.10 平均意义下的触点点云	75
图 4.11 不同被试的触点点云	75
图 4.12 不同语言模型下解码器的预测准确率	77
图 4.13 通用化模型和个性化模型的预测准确率	78
图 4.14 打字速度	80
图 4.15 错误率	81
图 4.16 单词解码器的准确率	82
图 4.17 文本输入时间构成	83
图 4.18 选词时间	83
图 4.19 主观评分（1 分最差，7 分最好）	84
图 4.20 比较智能打字指环与智能手指食指打字	85
图 5.1 TypeBoard 原理介绍	90
图 5.2 实验一的实验设置	96
图 5.3 实验任务图示	97
图 5.4 标注程序图示	97
图 5.5 有意性易混淆的案例	102
图 5.6 实验二的实验设置	103
图 5.7 所有的误触类型图示	105
图 5.8 用户差异和任务差异图示	107
图 5.9 采样时间和信道对误触识别准确率的影响	108
图 5.10 实验三的实验设置	111
图 5.11 三种键盘下被试打字速度对比	112
图 5.12 打字错误率图示	113
图 5.13 时间构成	114
图 5.14 两种键盘的触摸点云图示	115
图 5.15 主观评分	116
表 2.1 拟合触摸运动方程时对未知参数的约束	19

表 2.2 本实验中触摸交互数据的统计值，括号中的数值为标准差。时间单位为毫秒，距离单位为厘米，力的单位为克，速度单位为米每秒。	23
表 2.3 触摸运动模型与第 3、4、5 章所述触摸交互技术之间的联系。	27
表 3.1 触摸姿态分类的准确率（括号中数值为标准差）	42
表 3.2 指环佩戴在不同位置上时的触摸检测准确率（触摸检测延迟为 20 毫秒）	
46	
表 3.3 表格对比了不同触摸检测技术的性能，括号中的数值为标准差。	53
表 4.1 触摸手势识别技术在不同无源表面上的表现	71
表 4.2 表格展示了智能打字指环与相关技术的对比。请注意，这一对比仅供参考，由于不同文献中的实验方法和实验条件不同，此对比不是完全公平的。	86
表 5.1 表格中展示了三个阶段中喂入支持向量机模型的特征组。除了“时空相关性特征”以外，其余特征组都需要提取 100 毫秒内的时域特征，包括最大值、最小值、平均值、峰度和偏度。	98
表 5.2 阶段二防误触算法的预测失败案例，N 表示失败案例计数。	101
表 5.3 实验一和实验二之间的用户打字行为差异。本节用 T 检验来评估差异的显著性，如果 Levene 测试拒绝了数据的同方差性，本节将使用不等方差 T 检验来代替 T 检验。	105

符号和缩略语说明

MR	混合现实 (Mixed Reality)
IMU	惯性传感单元 (Inertial Measurement Unit)
OOV	词库外 (Out of Vocabulary)
RGB	红绿蓝三色标准 (Red Blue Green)
RMSE	均方根误差 (Root Mean Squared Error)
SD	标准差 (Standard Deviation)
SVM	支持向量机 (Support Vector Machines)
Qwerty	英文 26 键全键盘
KNN	邻近算法 (K-Nearest Neighbor)
RQ	研究问题 (Research Question)
UER	未纠正错误率 (Uncorrected Error Rate)
CER	已纠正错误率 (Corrected Error Rate)
WPM	每分钟输入英文词数 (Word Per Minute)
CPM	每分钟输入中文字数 (Character Per Minute)

第1章 引言

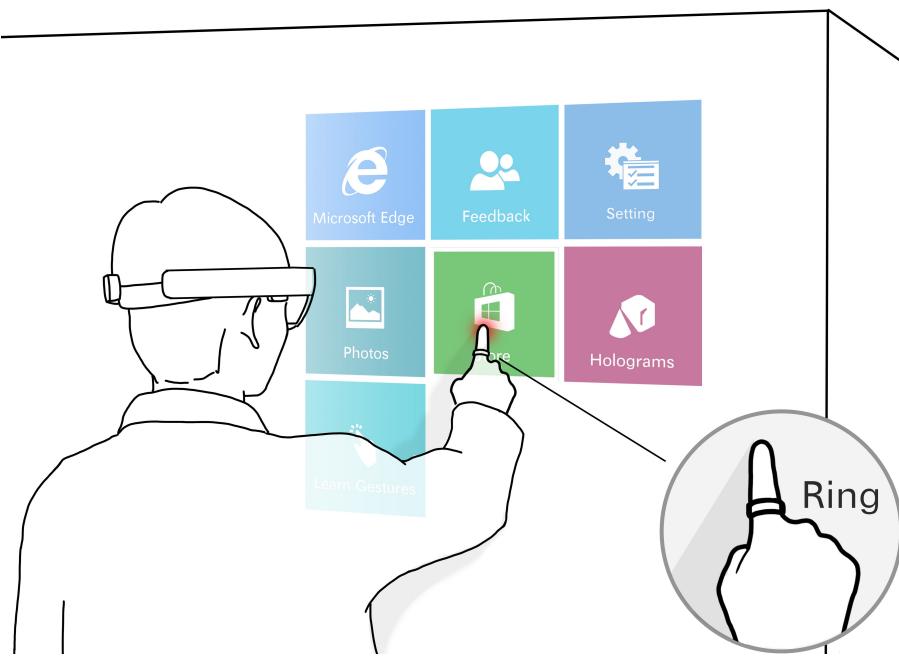
触摸交互是自然人机交互的重要组成部分，是人主动控制手指触摸交互表面，通过点击、长按、滑动等手势向计算机输入信息的方式。触摸屏作为触摸交互的主要载体已问世数十年，然而，触摸交互在普适性、响应性和意图性上仍然存在很大的改进空间：为提供最佳用户体验，触摸交互应摆脱有源交互表面的束缚，拓展至无源的桌面、墙面上；触摸交互的端到端延迟应控制在 20 毫秒以内，以免人察觉到延迟；触摸交互的防误触能力应落实到人的意图层面，任何不表达输入意图的触摸都应被过滤。在此背景下，本文提出触摸运动模型，揭示触摸前后极短时间内手指的运动规律，即手指位移、速度和加速度的时空运动特征，对增强触摸交互的普适性、响应性和意图性具有指导意义。本章首先介绍触摸交互的重要性和优化空间，随后通过文献综述总结现有触摸交互技术及其不足，接下来简述触摸运动模型及其研究内容，最后介绍论文的组织结构。

1.1 选题背景及意义

1.1.1 触摸交互的重要性

触摸交互是当前最重要的人机交互方式之一。2021 年，手机、平板电脑等触摸屏设备的全球出货量达到 15.1 亿台，而笔记本、台式机等基于键鼠交互的设备的全球出货量为 3.6 亿台^[3]，其规模仅为触摸屏设备的 23.8%。相比于键鼠交互，触摸交互具有便捷、易学、自然等优势，使其适用于包括老人、儿童在内的更广泛用户群体。

在未来，触摸交互仍是人机交互的重要研究课题。头戴式混合现实设备（简称 MR 头盔，例如微软的 Hololens2^[4]）很可能取代手机成为下一代智能终端。MR 头盔利用深度摄像头扫描物理环境，将虚拟元素叠加渲染在物理实体之上。如图1.1所示，MR 头盔中一种有前景的交互方式是将虚拟的用户界面渲染在无源表面上，支持用户通过触摸与用户界面交互^[5]。其中，无源表面指的是无（电）能源的表面，在本文中，无源表面特指生活中常见的非数字化表面，如桌子表面、墙面、人的手掌等等。在此设想下，触摸交互将摆脱有源触摸屏的束缚，让人获得在无源表面上触摸输入的能力。与目前 MR 头盔中流行的空中手势交互相比，触摸交互具有更高的效率和精度、更佳的主观用户体验^[6]。然而，目前无源表面的触摸交互技术仍未成熟，需要进一步研究。



在未来的可穿戴场景下，人应能在普通的表面如桌面、墙面上通过触摸交互向计算机输入信息。

图 1.1 无源表面触摸交互的构想

1.1.2 触摸交互的改进空间

尽管触摸屏已经问世数十年，但是触摸交互技术仍然在普适性、响应性和意图性三方面存在很大的改进空间。人的交互体验是人机交互的重要研究课题，因此，本小节以最佳用户体验为导向，讨论触摸交互技术的改进空间：

(1) 普适性：普适计算是一个强调和环境融为一体的概念，主张计算设备朝小型化、可穿戴的方向发展，直至消失在人们的视线当中，人们能在任何时间、地点，以任何方式与数字世界交互。在人机交互往普适计算发展的过程中，触摸屏等有源输入设备可能是第一批消失的，在此背景下，如何支持无源表面（如桌面、墙面、手掌）的触摸交互，是值得研究的课题。

(2) 响应性：响应性包括触摸检测的准确性和延迟两方面。触摸检测的准确性是触摸交互的基础，是人能否顺利完成交互任务的关键；人对触摸交互延迟的感官需求极高，能察觉到低至 24 毫秒的延迟^[1]，高于 100 毫秒的延迟会让用户感到极度厌烦^[7]，且影响交互效率^[1,8]。因此，若要提供最佳的用户体验，必须将触摸检测准确率提高至 99% 以上，延迟降低至 20 毫秒以内。

(3) 意图性：自然触摸交互中人的有意触摸和无意触碰混杂，给防误触技术带来困难。最佳的防误触技术应从用户意图的层面出发，将误触定义为任何不表达交互意图的触摸，过滤所有不表达交互意图的触摸，只对表达交互意图的触摸做出相应。然而，目前的触摸屏只能过滤轻敲、边缘误触等特殊的误触情况，防误触

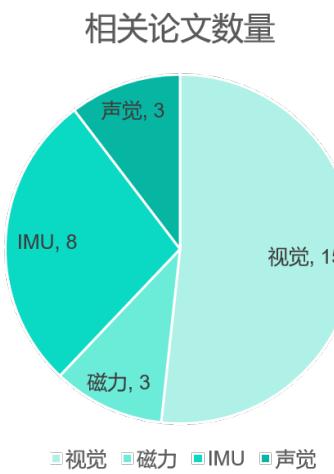
能力仍然需要改进。

电容触摸屏作为目前触摸交互的主要载体，在普适性、响应性和意图性上都远未达到优化目标，存在较大的改进空间。需要补充的是，在触摸交互技术的发展过程中，还曾存在**精准性**问题，即触摸 2D 位置的识别精度问题。然而，相比于鼠标交互，触摸交互是不精准的——人用手指能点中的最小目标直径为 4.3 毫米（点中概率为 95%）。造成触摸交互不精准的主要原因是胖手指问题^[9]，而非触摸交互技术不够精准，这使得继续优化触摸交互精准性对人机交互体验的意义不大。因此，触摸交互的精准性不在本文的讨论范围内。

1.1.3 为什么要建立触摸运动模型？

上一小节提到，触摸交互技术发展到今天，仍存在很大的改进空间。在此背景下，建立触摸运动模型能为触摸交互技术在普适性、响应性和意图性上的优化提供计算理论基础。其中，“运动”一词是本模型能够指导触摸交互技术的关键所在：

(1) 在普适性方面，触摸交互技术朝着无源表面触摸的方向发展。文献调研发现，近二十年来共有 29 篇关于无源表面触摸交互的高水平论文，它们全部依赖位移、速度和加速度等运动信号。其中，15 篇论文利用基于视觉的位移传感^[10-24]，3 篇论文利用基于磁力的位移传感^[25-27]，8 篇论文利用到基于惯性传感器的加速度传感^[2,5,28-33]，3 篇论文利用到基于麦克风的震动传感^[25-27]。以上观察说明，触摸交互技术若要摆脱触摸屏的束缚，触摸运动的规律是绕不开的话题。然而，上述论文始终缺少一个底层的模型作为计算理论基础。



近二十年 29 篇关于无源表面触摸交互的论文全部依赖运动信号。其中，视觉和磁力方法代表位移传感，IMU（惯性传感单元）和麦克风代表加速度和震动传感。

图 1.2 无源表面触摸交互的相关论文

(2) 在响应性方面，只有建立起触摸运动模型，触摸检测的准确率和延迟优化

才能有理可依。本文的后续分析表明，触摸交互的准确性和延迟优化并非独立的优化问题，而是此消彼长的关系：当一个触摸检测系统允许有更长的识别延迟时，它能用于检测触摸的信息就更丰富，准确率的上限就越高；但相应地，过长的识别延迟会直接影响到人的触摸交互体验。因此，触摸检测的准确性和延迟优化是需要权衡的问题，而触摸运动模型可以在两者的动态平衡中找到最有益于交互体验的解。

(3) 在意图性方面，运动信号是最能表征人的触摸意图的信号之一。人在组织一次触摸时，其心理是对手指的运动过程做出规划，并向运动系统发出指令。本文的后续分析表明，基于最优控制理论的触摸运动模型合理地描述了人的触摸心理，其拟合参数也能作为触摸有意性的判断依据。

综上所述，触摸运动模型能刻画人的触摸运动规律，剖析人的触摸心理，从而指导触摸交互技术在普适性、响应性和意图性上的优化。

1.2 研究现状

1.2.1 触摸交互技术

触摸交互是手机、平板电脑等手持智能终端上最常用的人机交互方式。大部分触摸交互设备依赖有源的触摸屏技术，需要在交互表面上安装传感器，按技术类型可分为电阻式^[34-35]、电容式^[36-37]、光学^[38-40]和声学^[41-42]触摸屏。电容式触摸屏是触摸交互的主要载体，其工作原理是利用低电压交流电场识别人手指与屏幕接触形成的耦合电容。然而，以上触摸屏感知技术要求在所有交互表面上安装传感单元，不可能在可接受的成本内支持普适计算中所展望的随时随地、随心所欲的触摸交互模态，因此，触摸屏交互技术面临普适性不强的问题。为提高触摸交互的普适性，部分研究者致力于将触摸交互从触摸屏的束缚中解放出来，实现无源表面上的触摸交互技术。

基于摄像头视觉的触摸交互技术是支撑无源表面上触摸交互的方法之一。视觉触摸交互技术可以将摄像头部署在有源触摸屏上^[38-40]，也可以将摄像头部署在场景空间中，或使用可穿戴的摄像头^[10-11]，从而支持无源表面触摸交互。相关文献中已经提出了几种用于无源表面触摸检测的视觉方案，包括激光雷达^[12]、RGB摄像头^[13-16]、红外摄像头^[17]和热成像摄像头^[18]等方案。近年来，深度摄像头变得越来越廉价，这引起了研究者们对基于深度摄像头的触摸交互技术的广泛研究，研究内容主要包括交互设计^[11,13,19]和精准性问题^[10-11,20-21]。然而，目前的视觉触摸交互技术存在响应性不高的问题，甚至难以准确判断手指是否接触了交互表面^[11,19]。已知文献中的视觉触摸交互技术都使用阈值方法传感触摸事件^[10-11,13,19,21-22]。例

如，在代表性工作 MRTouch^[11]中，研究者利用 Hololens 混合现实头盔的前置摄像头判断触摸事件的算法流程是：当手指指尖和交互表面之间的距离降至 10 毫米以内，汇报触摸事件发生；在手指接触交互表面以后，若手指远离交互表面，距离超过了 15 毫米，则汇报触摸事件结束。MRTouch 作为混合现实头盔上最高水平的无源表面触摸交互技术，其未识别率为 3.5%，误触率高达 19.0%，端到端延迟高达 180 毫秒，这说明视觉触摸交互技术仍需改进。造成上述问题的原因有两点：首先，摄像头对手指和交互表面位置的识别存在误差，特别是对缺少视觉特征点的纯色表面，识别精度更低；其次，从摄像头的位置看去，手指与交互表面的接触位置通常被遮挡，特别是在最有应用前景的混合现实头盔情景下，更是一定会被遮挡。

基于震动的触摸交互技术是支撑无源表面触摸交互的另一种方法，其工作原理是利用惯性传感器或麦克风感知触摸产生的震动和声音，既可以将传感器部署在有源触摸屏上^[41-45]，也可以将传感器部署在人的手指^[5,28-29]或手腕^[30]上，从而支持无源表面触摸交互。过去，大部分相关工作致力于提高触摸交互的精准性，即提高对手指位置的追踪精度^[2,31]，还有的工作致力于识别用户用哪跟手指进行的触摸^[46]，或者识别交互表面的材质^[47]。然而，较早前工作忽视了触摸交互的响应性问题，仅利用阈值方法来识别触摸事件的发生^[2,31-32]，准确率不超过 89.8%，端到端延迟不低于 50 毫秒。直到本文作者在 2019 年提出基于惯性传感指环的低延迟触摸检测技术^[5]，基于震动的无源表面触摸交互技术才在响应性上达到用户无法察觉到延迟的标准，后来有其它研究者进一步改进了这项技术，使其适用于更多触摸手势^[28]和腕带式传感器^[30]。

从以上文献能看出，触摸交互技术分为有源触摸屏技术和无源表面触摸交互技术两大类。其中，无源表面触摸交互技术作为未来混合现实场景的核心技术，有两种技术实现方案，一是基于摄像头的视觉方法，二是基于手指触摸瞬间震动的方法。视觉和震动方法的本质都是利用了手指触摸过程的运动规律，即位移和加速度的规律。虽然现有工作通过阈值方法、规则或机器学习利用到了触摸的运动规律，但还没有工作系统地将规律总结起来。在此背景下，本文提出了基于最优控制理论的触摸运动模型，揭示触摸前后极短时间内手指的运动规律。

1.2.2 触摸交互模型

在对触摸交互建模的相关文献中，大部分工作致力于建模触摸的指点过程，触摸指点是人的手指从触摸屏上的初始点出发，移动到另一个目标点的运动过程。从时间轴上看，触摸指点过程以目标点的出现为起点，以手指点中目标点为结束。触摸指点建模的代表性工作是利用费茨定律^[48]来预测标准触摸界面中手指触摸指点过程的移动时间^[49]，通过为特定触摸屏设备构建费茨模型，可以在给定初始点

位置和目标点位置的情况下预测手指移动的时间。Wobbrock 等人通过费茨定律的拓展来预测触摸指点过程的精度^[50]，该模型在已知运动时间的情况下预测触摸指点的精度。

尽管大量的研究工作致力于建模触摸指点过程^[51-55]，但是研究者们对触摸运动模型的理解还很浅。触摸运动模型指的是手指接触交互表面前后极短时间内手指的运动规律，包括手指的位移、速度和加速度规律。触摸运动模型对触摸交互技术的普适性、响应性和意图性具有指导意义。据本文作者所知，Xia 等人的论文“零延迟触摸技术”^[56]是先前工作中与触摸运动模型最为相关的文献，他们提出用抛物线拟合手指触摸前的运动轨迹，可用于预测手指的触摸着陆位置和触摸时间。这份工作对触摸运动模型的探索是初步性的：首先，论文仅选取了单一的触摸任务（大屏幕上大跨度的触摸指点任务）作为研究对象，该任务下手指运动幅度大，其手指运动规律与日常生活中常用的轻触差异很大。例如，论文声称能在手指与交互表面相距 3.2 厘米时准确预测触摸事件，但在大多数便携设备的自然触摸中，手指不会离开交互表面超过 2 厘米，这说明该论文在触摸交互的任务选择上有很大的局限性。第二，这份工作以抛物线（二次方程）拟合手指的运动轨迹，公式较为简单，不能为触摸交互技术提供大量有用的信息，作者也无法从手指的生理机制层面解释方程的意义。

相比于触摸前后的手指运动规律，人们对手部三维运动规律的认知更多。生物机械科学家^[57]和神经系统科学家^[58-59]热衷于研究人的手部三维运动，他们的研究兴趣主要集中在了解各种运动学特征，例如肌肉驱动、关节扭矩和手部运动过程中人的认知规划，等等。Flash 等人基于最优控制理论，通过定义目标函数和运行优化算法建立了无约束的端到端手臂运动模型^[58]，他们发现对手臂运动急动度进行最优化会生成可接受的轨迹方程。遵循类似的方法，Uno 等人通过对手关节扭矩进行最优化生成了另一个可接受的轨迹方程^[57]。本文所介绍的基于最优控制理论的触摸运动模型与 Flash 等人的工作较为类似，不同点在于：（1）Flash 等人的工作研究无约束的端到端手臂运动，而本文所介绍的触摸运动，本质上是有约束（交互表面阻挡）的端到端手指运动；（2）Flash 等人的工作只考虑模型的数学表达，而本文除了考虑模型的数学表达，还探讨了模型的计算方法，力求触摸运动的数学模型在计算上是可用的。

1.3 研究内容

触摸交互是人主动控制手指触摸交互表面，通过点击、长按、滑动、拖拽等手势向计算机输入信息的方式。目前，触摸交互技术的改进难点在于触摸瞬间的

识别是否具有强普适性^[5,10-11]、高响应性^[8,56,60]和准确的有意性^[61-63]，这要求研究者在触摸前后极短的时间跨度上优化触摸交互技术。在此背景下，本文提出了基于最优控制理论的触摸运动模型，揭示了触摸前后极短时间内的手指运动规律。围绕触摸运动模型及其应用，本文提出以下三点研究内容：

1.3.1 触摸运动模型的数学表达

触摸运动模型是描述一次触摸中，手指接触到交互表面前后极短时间内手指的运动规律。触摸运动模型的数学表达，也称触摸运动方程，是手指的位移随着时间变化的函数 $x(t)$ ，该函数对时间求导可以得到手指速度和时间的关系 $v(t)$ ，求二次导可以进一步得到手指加速度的时间的关系 $a(t)$ 。为了让函数 $x(t)$ 更贴近客观事实，同时易于计算、能提供有用信息、能指导触摸交互技术发展，本文提出以下三点优化思路：

- **尊重原理：**在求解触摸运动模型的数学表达时，研究者应充分尊重人的认知和行为规律，而不能仅通过实验数据进行拟合。举例来说，生物机械科学揭示了人的运动能力（如最大速度、加速度和急动度）存在上限^[64]；神经系统科学揭示了人在组织一次肢体的运动过程时，人在潜意识下存在一个全局的最小化目标，或称“代价”，如肢体运动的急动度^[58]或关节的扭矩^[57]。尊重上述先验知识将让模型表达式的求解过程事半功倍，也能确保模型更贴近客观事实。
- **拟合良好：**触摸运动模型的数学表达需要良好地拟合实验测量结果。触摸交互的适用设备、用户群体和交互任务都是多样化的，研究者在采集触摸的实验数据时，需要排除实验设备造成的差异，覆盖广泛的用户群体和交互任务。触摸运动模型应适用于描述任何触摸交互行为，且拟合良好，而不能仅针对特定交互任务和人群进行拟合。
- **简洁明了：**在尊重原理和拟合良好的前提下，触摸运动模型的数学表达应尽量简洁明了，可以选择性地舍弃对拟合精度影响较小的变量，这将有利于触摸运动模型的计算和应用。经典的人机交互模型往往是简洁的，例如，尽管费茨定律在其诞生后经历过一些复杂化的修补^[49-50,65-66]，但原版费茨定律仍然是人机交互领域中应用最广的理论之一^[48]。

1.3.2 触摸运动模型的计算方法

触摸运动的计算方法是已知触摸运动方程的表达式，利用位移、速度和加速度等传感信号拟合触摸运动方程参数的计算方法，包括两部分：（1）信号的处理和降噪；（2）通过信号拟合触摸运动方程。为了提高触摸运动模型的计算方法的

性能，本文提出以下三点优化思路：

- **信号降噪：**原始的传感信号往往伴随着噪声，在拟合触摸运动方程之前，应将信号的噪声降至较低水平。对于单一的信号源，可利用信号的时序关系降低噪声；特别地，当存在多模态信号时，可利用信道之间的互信息降低各信道的噪声。
- **拟合良好：**触摸运动模型的计算方法是针对一次特定的触摸，求解其触摸运动方程参数的方法，解得的方程应尽可能贴近传感器测量结果。更具体地，一般的运动传感信号，如基于视觉的位移信号和基于惯性传感器的加速度信号，其误差服从正态分布，在此情况下应采用最小二乘法进行拟合。
- **高效计算：**人在触摸交互时对系统响应性的要求极高，能感知到低至 24 毫秒的端到端延迟^[1]，在此背景下，触摸运动模型若要在实时的触摸交互技术中发挥作用，就需要有一套毫秒级的计算方法。

1.3.3 基于触摸运动模型的交互优化技术

触摸运动模型揭示了触摸事件前后极短时间内手指运动的规律，与触摸交互技术息息相关，为提高触摸交互技术的普适性、响应性和意图性提供了计算理论基础。然而，触摸运动模型作为一种理论，在实际应用的过程中可能还会遇到一些困难，因此，应该如何利用触摸运动模型优化触摸交互技术，也是本文重要的研究内容。本文选取了三种典型的触摸交互技术，来探索利用触摸运动模型优化交互技术的可行性：

- **触摸检测技术：**触摸检测是识别触摸是否发生的技术。良好的触摸检测技术能够高响应性地识别别人的手指是否接触交互表面，其中，高响应性包括高识别准确率和低延迟。
- **触摸手势识别技术：**触摸手势识别是区分点击、抬起、长按、滑动等触摸手势的技术。在触摸检测的基础上，良好的触摸交互系统应能准确区分常用的触摸手势，以支持更丰富的交互语义。
- **触摸意图识别技术：**触摸意图识别是区分有意触摸和无意触碰的技术。良好的触摸意图识别技术应从用户的交互意图层面出发，准确区分表达交互意图的有意触摸，和不表达任何交互意图的误触。

为了验证模型对触摸交互技术的优化效果，本文选取了两种典型的触摸交互作为评测实验的实验任务，分别是：(1) **触摸指点任务**，即人通过触摸点中交互界面中目标点的任务，指点任务是最基础的人机交互任务之一；(2) **触摸文本输入任务**，即基于触摸交互的文本输入任务，文本输入作为最快速、最复杂的触摸交互任务之一，被列入本文的评估范围。

1.4 主要研究成果

本文提出了基于最优控制理论的触摸运动模型，揭示了触摸前后极短时间内手指的运动规律，探讨了如何利用运动传感信号（包括位移、速度和加速度）拟合触摸运动方程。本文详细描述了触摸运动模型（第2章），基于触摸运动模型，本文针对经典的触摸交互任务（目标选择和文本输入），优化了触摸检测技术（第3章）、触摸手势识别及文本输入技术（第4章）和触摸防误触技术（第5章），提升触摸交互的效率和用户体验。

（1）提出了基于最优控制理论的触摸运动模型

实验观察和验证发现，人产生触摸意图时，对触摸运动做出的规划可描述为：“在规定时间 t_1 内，最平稳地将手指从初始点 x_0 移动到目标点 x_1 。”其中，目标点 x_1 是位于交互表面之下的虚构点。在触碰到交互表面之前，手指会沿着最平稳的时空轨迹向目标点 x_1 移动，直到接触交互表面而停止。其中，“最平稳的”指的是最小化手指运动急动度的平方的积分^[58]，由此约束可解得触摸运动的方程如下所示：

$$x(\tau) = \begin{cases} x_0 + (x_1 - x_0)(6\tau^5 - 15\tau^4 + 10\tau^3) & \tau \leq \tau_c \\ 0 & \tau > \tau_c \end{cases} \quad (1.1)$$

其中， $\tau = \frac{t}{t_1}$ 是运动的时间进度， τ_c 表示手指触摸到交互表面时的时间进度。以上就是触摸运动模型的数学表达。为了将模型应用在触摸交互技术中，本文还讨论了触摸运动模型的计算方法，即利用位移、速度和加速度等运动传感信号拟合触摸运动方程的算法。最后，本文通过用户实验证明触摸运动模型符合客观物理事实、能够良好拟合实验数据、显著优于现有模型。

触摸运动模型对触摸交互技术具有指导意义，应用方向包括：(1) 触摸检测；(2) 触摸手势识别；(3) 触摸意图推理。基于触摸运动模型，本文对这三项技术进行深入优化，做出后续三点主要贡献：

（2）指环上的高准确低延迟触摸检测技术

针对现有触摸检测技术普适性不强，响应性不高的问题，本文提出了指环上的高准确低延迟触摸检测技术。目前主流的电容触摸屏技术存在普适性不强的缺点，用户只能在专用触摸屏设备上进行触摸交互，不符合普适计算中随时、随地、随心交互的原则。基于惯性传感指环的触摸检测技术使用户能在无源物体表面（如桌面、墙面）上触摸输入^[2,31]，增强了触摸交互的普适性。然而，先前基于惯性传感指环的技术采用阈值方法判断触摸事件，延迟高达 200 毫秒，准确率仅为 85%，不满足触摸交互的低延迟需求。

基于触摸运动模型，本文提出了指环上的高准确低延迟触摸检测技术。实验结果显示，本技术的触摸检测延迟低至 10 毫秒，准确率超过 99%。低延迟该技术的关键词，由于人在触摸交互中无法察觉到低于 24 毫秒的延迟，本技术为用户提供了最佳用户体验，与先前技术相比实现了质的飞越。

(3) 指环上的触摸手势识别与打字技术

针对新型显示设备（例如混合现实头盔和智能电视）上缺少高效文本输入法的问题，本文提出了基于惯性传感指环的无源表面文本输入方法，佩戴指环的用户能在普通的物体表面（如桌面、墙面）上快速打字。

触摸手势识别技术是触摸文本输入的前提条件。本文结合触摸运动模型，实现了无源表面上基于惯性传感指环的触摸手势识别技术，能准确识别包括手指触摸、抬起、长按和滑动在内的多种触摸交互手势。在触摸手势集的基础上，本文提出了无源表面文本输入方法。实验结果显示，用户打字速度为每分钟输入 20.6 个英文单词，达到手机打字效率的 86.5%。由于该文本输入方法与显示设备解耦，它为混合现实、智能电视等多种场景提供了有效的文本输入方案，具有很强的普适性。鉴于文本输入是最快的、复杂的触摸输入任务之一，该文本输入方法的成功也证明了低延迟触摸检测技术和触摸手势识别技术的实用性。

(4) 平板电脑上的触摸意图推理键盘

针对现有平板电脑上，连续触摸输入误触频发的问题，本文提出了基于触摸运动模型的防误触技术。平板电脑文本输入是典型的连续触摸任务，是最复杂的触摸交互任务之一。目前，用户在触摸屏上十指打字时需要将手指悬空，以避免误触发生，久而久之会产生疲劳的问题；而若用户不将手指悬空，则手指在贴近触摸屏时会引发误触。

解决上述问题最直接的方法是开发一款强力的防误触技术，让用户在平板电脑十指打字时可以将非交互手指休息在屏幕上，而不引发误触，系统只对表达打字意图的触摸做出响应。通过实验观察和数据分析，作者发现：(1) 同样是手指触摸屏幕，有意触摸和误触在触摸运动方程的参数上是有差异的，触摸的速度、力度、时长都是触摸有意性的重要表征；(2) 有意触摸中，手指触摸运动方程的参数与其它手指独立，而多指休息导致的误触中，不同手指的触摸运动方程的参数具有强相关性。结合以上规律，作者提出了面向连续触摸输入的防误触技术，识别准确率达到 99%，且允许用户在十指打字时将非交互手指休息在触摸屏上，而不会引发误触。实验表明，该技术改变了用户的打字行为，用户将手指休息在触摸屏上，打字行为更自然、更抗疲劳，打字速度也提升了 20%。

1.5 论文组织结构

本文后续章节的组织结构如下：

第2章，介绍基于最优控制理论的触摸运动模型，包括触摸运动模型的数学表达、计算方法、实验评估和指导意义。该章所介绍的触摸运动模型是全文的核心，是后续三章交互优化技术的计算理论基础。

第3章，介绍指环上的高准确低延迟触摸检测技术，包括指环触摸交互的自然性研究、基于机器学习的触摸检测技术和基于触摸运动模型的触摸检测技术。

第4章，介绍指环上的触摸手势识别和打字技术，包括指环上的触摸手势识别、智能指环打字技术的交互设计、智能打字指环的解码器设计和打字技术评测。

第5章，介绍平板电脑上的触摸意图推理键盘，包括基于触摸运动模型的防误触技术介绍、迭代了两轮的防误触键盘算法设计和防误触键盘评测。

第6章，对本文内容进行总结和展望。

本文的核心是基于最优控制理论的触摸运动模型，对该模型的介绍主要在第2章。在模型的指导下，本文的第3、4、5章介绍了三种基本的触摸交互技术。这三章内容独立成章，会具体介绍技术的背景、交互设计、实验评测等等，而触摸运动模型对这些技术的优化主要集中在章节3.5.1、4.4.2和5.3.2。

第2章 基于最优控制理论的触摸运动模型

最优控制理论是数学最优化的分支，研究使动力控制系统的性能指标实现最优化的综合方法。自1962年Pontryagin等人提出最优控制理论以来^[67]，该理论被广泛应用于空间技术^[68]、运筹学^[69]、经济调控^[70-71]等重要领域。1985年，Flash等人指出，最优控制理论还可用于描述人的手部运动过程：人在组织手部运动时，大脑不会具体地控制肩关节和肘关节的旋转角度，而是将注意力集中在手掌，试图在规定时间内，最平稳地将手部从初始点移动到目标点。其中，“最平稳地”指的是人下意识地最小化手部运动急动度平方的积分：

$$\min \int_0^{t_1} \left(\left(\frac{d^3x}{dt^3} \right)^2 + \left(\frac{d^3y}{dt^3} \right)^2 + \left(\frac{d^3z}{dt^3} \right)^2 \right) dt \quad (2.1)$$

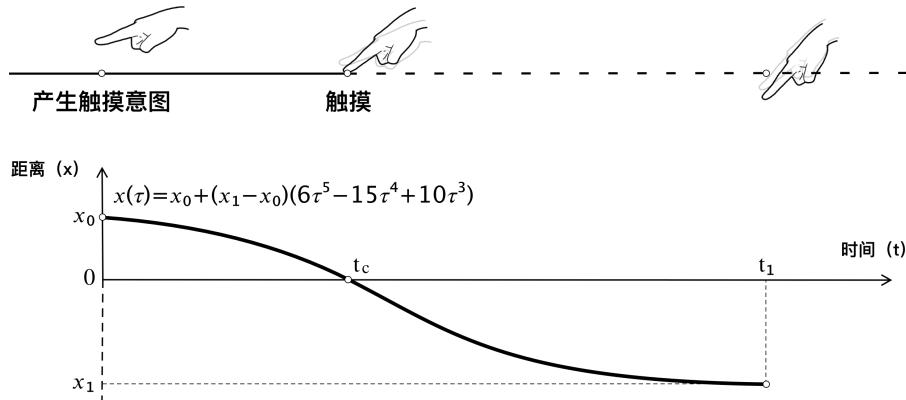
其中， $x(t)$ ， $y(t)$ 和 $z(t)$ 是手部位置在直角坐标系三个轴上的分量随时间变化的函数。受到这份工作的启发，本文作者通过实验验证发现，公式2.1约束下的运动方程同时也是触摸运动的良好拟合。因此，本文提出了基于最优控制理论的触摸运动模型。本章分四小节介绍触摸运动模型：

1. **触摸运动模型的数学表达：**提出描述触摸运动的数学方程，描述触摸运动中手指位移与时间之间的关系。
2. **触摸运动模型的计算方法：**提出利用运动传感信号拟合触摸运动方程的计算方法，其中运动传感信号包括手指的位移、速度和加速度。
3. **触摸运动模型的实验评估：**通过实验解释模型的提出过程，评测触摸运动模型的拟合精度。
4. **触摸运动模型的指导意义：**讨论触摸运动模型对几种基础触摸交互技术的指导意义，介绍模型对触摸检测、触摸意图推理和触摸手势识别的优化原理。

2.1 触摸运动模型的数学表达

触摸运动模型是描述触摸前后极短时间内手指运动规律的模型，其数学表达是触摸运动中手指位置 x 与时间 t 关系的函数。触摸运动模型的数学表达也称触摸运动方程。为简化触摸运动模型的表达和计算，手指位置 x 只考虑手指在交互表面之上的高度，而不考虑手指在交互表面上的2D坐标——如引言所述，目前触摸交互技术的瓶颈在于普适性、响应性和意图性，而不在于2D坐标的精准性。触

摸运动方程在时间轴上的起点是人产生触摸意图的瞬间 $t = 0$, 结束点是人在心理上认为的触摸动作结束的时间 $t = t_1$, 这段时间内包含了手指接触交互表面的瞬间 $t = t_c$ ($0 < t_c < t_1$)。一般情况下触摸运动的时长 t_1 介于 100 毫秒到 300 毫秒之间。



人组织一次触摸时, 假设交互表面凭空消失, 人会在规定时间 t_1 内, 最平稳地将手指从初始点 x_0 移动到目标点 x_1 , 其时空运动轨迹如图中的公式所示。

图 2.1 无约束运动模型

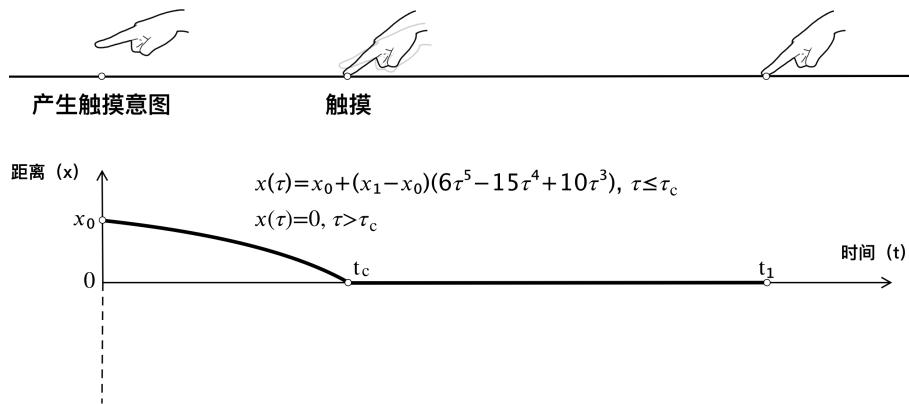
为方便理解, 此处引入入无约束运动模型, 该模型基于如图2.1所示的假想情形: “若在触摸瞬间交互表面凭空消失, 人会在规定时间 t_1 内, 最平稳地将手指从初始点 x_0 移动到目标点 x_1 ”。其中, “最平稳地”指的是最优控制理论对肢体运动的约束 (公式2.1), 在此约束下解得无约束运动模型的数学表达为:

$$x(\tau) = x_0 + (x_1 - x_0)(6\tau^5 - 15\tau^4 + 10\tau^3) \quad (2.2)$$

其中, $\tau = \frac{t}{t_1}$ 是触摸运动时间进度。上述公式中 $x_0 > 0$ 是容易理解的, 即人在产生触摸意图时手指位于交互表面的上方。但值得注意的是, 公式中有 $x_1 < 0$, 其含义是: “人在组织一次触摸运动时, 其心理并非将手指带到交互表面上, 而是将手指带到交互表面之下一个虚构点上”。

然而, 如图2.2所示, 在实际情况下, 交互表面的存在中断了上述运动过程, 手指在接触到交互表面时瞬间停止了。在手指接触到交互表面之前 ($t < t_c$), 手指运动的时空轨迹与无约束运动模型 (公式2.2) 一致; 而在手指接触到交互表面之后 ($t > t_c$), 手指静止在交互表面上。因此, 触摸运动模型的表达式应为下述分段函数:

$$x(\tau) = \begin{cases} x_0 + (x_1 - x_0)(6\tau^5 - 15\tau^4 + 10\tau^3) & \tau \leq \tau_c \\ 0 & \tau > \tau_c \end{cases} \quad (2.3)$$



一次触摸运动中，人的心理是在规定时间 t_1 内，最平稳的将手指从初始点 x_0 移动到目标点 x_1 ，但交互表面的阻挡使得手指瞬间停止。触摸运动的时空运动轨迹如图中的公式所示。

图 2.2 触摸运动模型

通过函数对时间进度 τ 的求导不难发现，触摸运动模型同时揭示了手指速度 v 和加速度 a 的变化规律：

$$v(\tau) = \begin{cases} (x_1 - x_0)(30\tau^4 - 60\tau^3 + 30\tau^2) & \tau < \tau_c \\ 0 & \tau > \tau_c \end{cases} \quad (2.4)$$

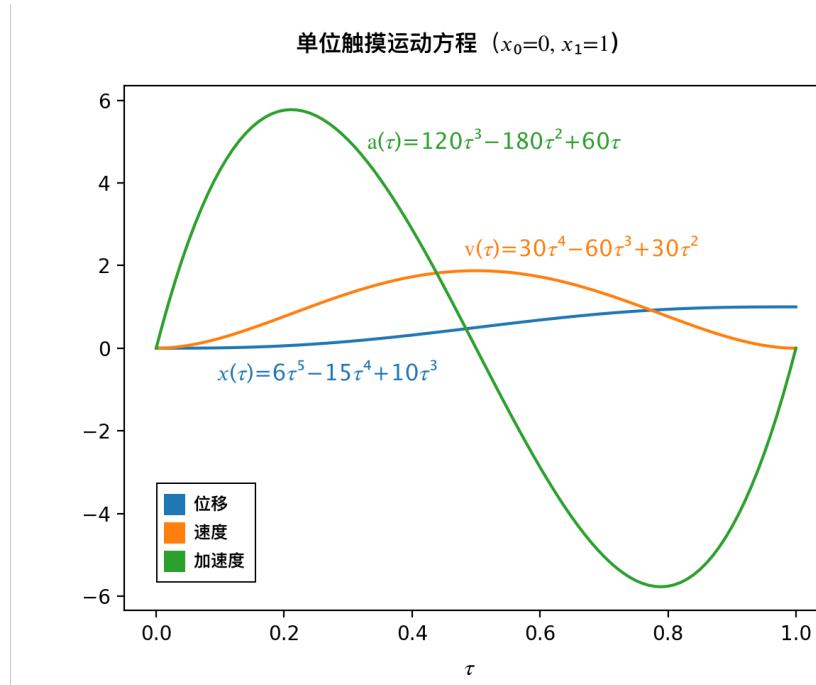
$$a(\tau) = \begin{cases} (x_1 - x_0)(120\tau^3 - 180\tau^2 + 60\tau) & \tau < \tau_c \\ 0 & \tau > \tau_c \end{cases} \quad (2.5)$$

综合以上三个公式中可以看出，触摸运动过程分为两部分，前半部分中手指处于一个向下运动的过程，其运动的时空轨迹较为复杂，是本节讨论的重点；后半部分手指静止在交互表面上，手指的位移、速度和加速度恒为零。本节的剩余内容主要讨论触摸运动方程的前半部分，也就是触摸运动的向下过程。

2.1.1 单位触摸运动方程

描述一次触摸运动的向下过程只需要三个参数，分别是 x_0 、 x_1 和 t_1 。其中， x_0 是人产生触摸意图时手指的位置， x_1 是人组织这次触摸运动时假想的目标点， t_1 是人组织触摸运动时假想的运动时长。触摸运动轨迹的拟合过程，就是求解 x_0 、 x_1 和 t_1 的过程。

如图2.3所示是单位触摸运动方程，是归一化的触摸运动时空轨迹 ($x_0 = 0, x_1 = 1, \tau = \frac{t}{t_1}$)。由于拟合触摸运动方程只涉及 x_0 、 x_1 和 t_1 这三个参数的求解，所以任何触摸运动方程的都是单位触摸运动方程在 X 轴和 Y 轴上分别缩放和平移的结果。



单位触摸运动方程是归一化的触摸运动时空轨迹，其位移、速度和加速度的轨迹如图所示。

图 2.3 单位触摸运动方程

果。

2.1.2 触摸运动方程的推导过程

本小节将介绍触摸运动方程的推导过程。与许多经典模型一样（如费茨定理^[48]），基于最优控制理论的触摸运动模型建立在一些必要的前提条件（假设或简化）之上，以保证模型简洁明了，这将有利于模型的计算和应用。触摸运动方程的手指向下的过程有两个前提条件：

(1) 触摸运动中手指的向下运动过程符合无约束运动模型：若交互表面凭空消失，人会在规定时间 t_1 内，将手指从初始点 x_0 移动至虚构的目标点 x_1 ，且移动过程最小化了手指急动度的平方的积分（公式2.1）。

(2) 手指在初始点 x_0 和目标点 x_1 上静止：在这两点上手指的速度和加速度为零。

前提一是一个假设，该假设来自 Flash 等人对手部肢体运动的建模^[58]。一方面，该假设已经在先前工作中得到了充分的验证，证明其对肢体运动的拟合良好；另一方面，在这一假设下本文提出了触摸运动方程，也良好地拟合了触摸运动的实验数据。

前提二是本文对触摸运动规律的有意简化。事实上， x 在触摸运动的初始点和

目标点上仅仅是接近于静止，而非完全静止。特别是，在初始点上手指可能有初速度；在目标点上手指可能有一个向上的加速度。尽管与事实有一定的偏差，本模型仍然保留了这一简化：一方面，该简化对拟合程度的影响较小；另一方面，该简化将触摸运动方程的未知参数限制在三个以内（ x_0 、 x_1 和 t_1 ），提高了模型的可计算性和实用性。综上所述，触摸运动中手指的向下运动过程等价于以下最优化问题：

$$\begin{cases} x(t) \\ s.t. x(0) = x_0, x'(0) = x''(0) = 0, x(t_1) = x_1, x'(t_1) = x''(t_1) = 0 \\ \min \frac{1}{2} \int_0^{t_1} \left(\frac{d^3 x}{dt^3} \right)^2 dt \end{cases} \quad (2.6)$$

读者可能注意到，手指运动应受到人的运动能力的限制，比如，手指运动存在速度和加速度的上限。然而，上述最优化问题未包含对运动能力作出条件约束，这是因为，该最优化问题的解恰巧不会超出手部运动能力的限制。求解上述最优化问题即可得到无约束运动模型的数学表达（公式2.2），结合交互表面的存在中断了手指的无约束运动，进一步可推得完整的触摸运动方程（公式2.3）。

2.1.3 更复杂的触摸运动方程

上一小节提到，本文的触摸运动模型经过了有意简化。那么，如果没有这些简化，是否仍能求解出触摸运动模型的数学表达呢？答案是肯定的，只不过触摸运动方程会变得复杂，不利于模型的计算和应用。本小节介绍一些更复杂的触摸运动方程。

(1) 若手指仅在初始点 x_0 上静止，而在目标点 x_1 上存在一个向上的加速度 $a_1 (a_1 > 0)$ 。则触摸运动方程为，待拟合参数有 x_0 、 x_1 、 t_1 和 a_1 ：

$$x(\tau) = \begin{cases} x_0 + (x_1 - x_0)[(6 + \frac{a_1}{2})\tau^5 + (-15 - a_1)\tau^4 + (10 + \frac{a_1}{2})\tau^3] & \tau \leq \tau_c \\ 0 & \tau > \tau_c \end{cases} \quad (2.7)$$

(2) 若手指仅在目标点 x_1 上静止，而在初始点 x_0 上存在一个初速度 v_0 。则触摸运动方程为，待拟合参数有 x_0 、 x_1 、 t_1 和 v_0 ：

$$x(\tau) = \begin{cases} x_0 + (x_1 - x_0)[(6 - 3v_0)\tau^5 + (-15 + 8v_0)\tau^4 + (10 - 2v_0)\tau^3] & \tau \leq \tau_c \\ 0 & \tau > \tau_c \end{cases} \quad (2.8)$$

2.2 触摸运动模型的计算方法

触摸运动的计算方法指的是利用位移、速度、加速度等运动传感信号拟合触摸运动方程参数的计算方法。如图2.4所示，常见的针对手指的运动传感方法有（1）基于视觉方法的位移传感^[10-24]；（2）基于惯性传感器的加速度传感^[2,5,28-33]。由于大多数速度传感器利用位移除以时间来计算速度，相比于位移传感不提供额外的信息，因此本文不讨论速度信号。



左图是基于双目视觉的代表性位移传感器 Kinect^[72]，可捕捉人体四肢和手指的三维坐标；右图是惯性传感器，属于嵌入式设备，可传感自身的加速度和角速度。

图 2.4 常用的运动传感器

实验观察发现，触摸运动中手指向下运动过程的时长在 50 毫秒到 100 毫秒之间。因此，触摸运动的计算应利用触摸发生前 50 毫秒的数据拟合触摸运动方程。假设位移传感器的采样频率为 f_x ，加速度传感器的采样频率为 f_a ，则触摸发生前 50 毫秒内的测量数据为：

$$\begin{cases} [x_m(1), x_m(2), \dots, x_m(N_x)] & N_x = \lfloor 0.05f_x \rfloor \\ [a_m(1), a_m(2)s, \dots, a_m(N_a)] & N_a = \lfloor 0.05f_a \rfloor \end{cases} \quad (2.9)$$

触摸运动的计算方法可定义为利用上述测量数据拟合触摸运动方程（公式2.3）中参数 x_0 、 x_1 和的 t_1 算法，本节的剩余内容将介绍该算法。触摸运动的计算方法分为两个步骤：步骤一通过卡尔曼滤波融合各运动传感信号；步骤二通过最小二乘法拟合触摸运动方程。

2.2.1 卡尔曼滤波

在所有可能的在线滤波方法中，卡尔曼滤波对带有高斯白噪声的线性系统有着最佳的估计效果^[73]，对于位移、速度和加速度等运动信号而言，卡尔曼滤波是很好的预处理方法。如本章所述，针对手指的运动传感方法中最常见的是位移和加速度传感，因此本小节将介绍如何利用卡尔曼滤波平滑位移信号，同时估算速度。设位移信号的误差服从正态分布 $(0, \sigma_x^2)$ ，加速度信号的误差服从正态分布 $(0, \sigma_a^2)$ ，

设 $\mathbf{x} = \begin{bmatrix} x & \dot{x} \end{bmatrix}^T$, 则卡尔曼滤波的状态空间表达式为:

$$\mathbf{x}(k+1) = \mathbf{A}_d \mathbf{x}(k) + \mathbf{B}_d a_m(k) + \mathbf{u} \quad (2.10)$$

$$x_m(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}(k) + w = \mathbf{H} \mathbf{x}(k) + w \quad (2.11)$$

其中, $\mathbf{u} \sim (0, \mathbf{Q})$, $\mathbf{Q} = \begin{bmatrix} 0 & 0 \\ 0 & \sigma_a^2 \end{bmatrix}$, $w \sim (0, \sigma_x^2)$, $\mathbf{A}_d = \begin{bmatrix} 1 & T_a \\ 0 & 1 \end{bmatrix}$, $\mathbf{B}_d = \begin{bmatrix} T_a^2/2 \\ Ta \end{bmatrix}$, $T_a = \frac{1}{f_a}$ 。算得卡尔曼滤波的协方差矩阵 \mathbf{Q}_d 和 \mathbf{R}_d 为^[74]:

$$\mathbf{Q}_d = \sigma_a^2 \begin{bmatrix} T_a^3/3 & T_a^2/2 \\ T_a^2/2 & T_a \end{bmatrix} \quad (2.12)$$

$$\mathbf{R}_d = \frac{\sigma_x^2}{T_a} \quad (2.13)$$

由此卡尔曼滤波算法可以总结为:

(1) 时间更新:

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{A}_d \hat{\mathbf{x}}(k|k) + \mathbf{B}_d a_m(k) \quad (2.14)$$

$$\mathbf{P}(k+1|k) = \mathbf{A}_d \mathbf{P}(k|k) \mathbf{A}_d^T + \mathbf{Q}_d \quad (2.15)$$

(2) 测量更新:

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \mathbf{K}(k+1)[x_m(k+1) - \mathbf{H} \hat{\mathbf{x}}(k+1|k)] \quad (2.16)$$

$$\mathbf{P}(k+1|k+1) = [\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}] \mathbf{P}(k+1|k) \quad (2.17)$$

其中, 卡尔曼增益 $\mathbf{K}(k+1)$ 为:

$$\mathbf{K}(k+1) = \mathbf{P}(k+1|k) \mathbf{H}^T [\mathbf{H} \mathbf{P}(k+1|k) \mathbf{H}^T + \mathbf{R}_d]^{-1} \quad (2.18)$$

通过上述卡尔曼滤波即可估计第 k 帧的位移和速度 $\hat{\mathbf{x}}(k|k) = [x(k) \ v(k)]^T$ 。

若位移信号的采样间隔 T_d 和加速度的采样信号 T_a 不同, 且 $T_d/T_a = M$ 为整数, 则在每 M 帧同时执行时间更新和测量更新, 其余帧只执行时间更新:

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) = \mathbf{A}_d \hat{\mathbf{x}}(k|k) + \mathbf{B}_d a_m(k) \quad (2.19)$$

$$\mathbf{P}(k+1|k+1) = \mathbf{P}(k+1|k) = \mathbf{A}_d \mathbf{P}(k|k) \mathbf{A}_d^T + \mathbf{Q}_d \quad (2.20)$$

2.2.2 最小二乘拟合

拟合触摸运动方程 (公式2.3) 的过程是求解未知量 x_0 、 x_1 和 t_1 最大似然值的过程。本节提到, 拟合触摸运动方程需要用到最近 50 毫秒内的运动传感信号, 设近 50 毫秒内的第一帧数据 $(x_m(1), a_m(1))$ 对应触摸运动方程的时间戳 t_s , 即时间序列 $[x_m(1), x_m(2), \dots, x_m(N_x)]$ 和 $[a_m(1), a_m(2), \dots, a_m(N_a)]$ 对应触摸运动方程中 $t \in [t_s, t_s + 0.05]$ 的部分, 则 t_s 也是一个需要拟合的未知量。由于传感器误触符合正态分布, 应采用最小二乘法拟合触摸运动方程, 即求解以下最优化问题:

$$\begin{cases} x_0, x_1, t_1, t_s \\ \min \frac{\sum_{k=1}^{N_x} (x_m(k) - x(t_s + kT_x))^2}{\sigma_x^2} + \frac{\sum_{k=1}^{N_a} (a_m(k) - a(t_s + kT_a))^2}{\sigma_a^2} \end{cases} \quad (2.21)$$

在计算机算法中, 求解最优化问题的方法有很多^[75-77], 本章使用的最优化算法为信赖域方法^[78], 在 Python 的科学计算工具包 Scipy 中, 信赖域方法的名称是“trust-constr”。拟合时还需给未知参数设置合理的初始估计值和取值范围, 以提高拟合的效率和准确性。表格2.1展示了各未知参数的初始值估计和约束条件, 其中的距离单位为米, 时间单位为秒:

表 2.1 拟合触摸运动方程时对未知参数的约束

未知参数	初始值	下限	上限
x_0	0.01	$x_m(1)$	0.03
x_1	-0.035	-0.1	0
t_1	0.25	0.05	0.3
t_s	0.01	0	0.05

表格中各未知参数的初始值估计和取值范围均来自对大量触摸交互事件的观测。其中, 初始值估计来自大量触摸运动的均值, 例如, x_0 的初始值为 0.01 (米), 这是因为, 平均而言触摸交互发生之前人的手指在交互表面上方一厘米左右。同

理, t_1 的初始值为 0.2, 原因是触摸运动过程的时长约为 200 毫秒。表格中的上限和下限覆盖了大多数触摸交互的取值范围, 例如, x_0 的下限为 $x_m(1)$, 这是因为, 在触摸事件发生前的 50 毫秒时, 大部分情况下手指已经进入了向下运动的过程, 因此人产生触摸意图时手指与交互表面的距离 x_0 一定不小于 $x_m(1)$ 。

在实际应用中, 若仅有位移传感信号或加速度传感信号其一, 本小节介绍的计算方法也是可以运作的: 只需将公式2.21中待最优化函数中传感信号缺失的项置为零即可, 但拟合的准确性必然会受到影响。

2.3 触摸运动模型的实验评估

本节将介绍触摸运动模型的评估实验, 实验收集了人在几种典型的触摸交互任务中手指的运动信号, 其中, 触摸交互任务包括单次点击、连续快速点击、滑动手势、长按和拖拽。运动信号包括基于视觉方法的手指位移信号和基于惯性传感器的手指加速度信号。除了运动信号外, 实验还利用压敏触摸板收集了手指触摸之后向下的压力信号。根据上述数据, 本节将讨论触摸运动模型的提出过程、评测触摸运动模型的拟合精度。

2.3.1 数据采集实验

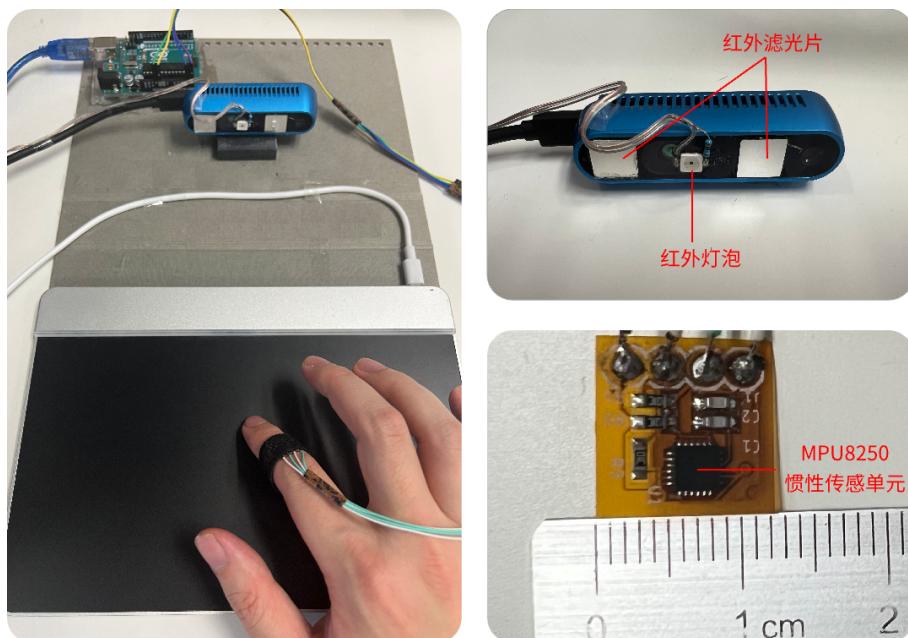
2.3.1.1 实验设计和过程

实验者从校园中招募了 12 名被试, 其中有 4 名女性, 被试的年龄从 19 岁到 25 岁不等, 平均年龄为 23.8 岁, 标准差为 2.27。所有的被试都是右撇子。如图2.5左侧所示是本实验的设置, 桌子上摆放着一块压敏触摸板, 作为实验的交互表面, 收集触摸的力度信息。被试的手指上被绑上了一个惯性传感器, 用于收集手指的加速度信号。在触摸板的远端有一个高速双目摄像头, 用于收集手指的位移信号。实验分为五个阶段, 在五种不同的触摸交互任务下收集被试的触摸运动信号, 交互任务包括:

(1) 单次点击: 被试在触摸板上用食指点击 50 次。被试以自己喜欢的方式以任意触点位置、点击力度和角度进行触摸, 但是被试需保证每两次点击之间有停顿, 先将手指移动至目标触摸位置上方, 瞄准之后再进行点击。实验共收集到 600 个单次点击数据片段。

(2) 连续快速点击: 被试在触摸板上用食指连续点击 50 次。在这一组实验中, 为了保证被试的触摸是连续且快速的, 实验者规定被试在一左一右的两个目标点上快速来回点击。实验共收集到 600 个连续快速点击数据片段。

(3) 滑动手势: 被试在触摸板上用食指滑动 50 次。被试在执行滑动手势时, 想



左图是本实验的设置，本实验收集被试在不同的触摸交互任务下的触摸运动信号。右图是实验设备的细节。

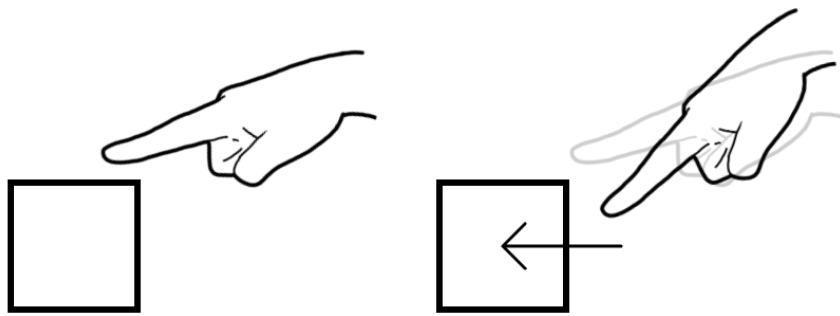
图 2.5 触摸运动模型评估实验设置

象自己在手机上做切屏动作，左滑和右滑交替进行。实验共收集到 300 次左滑数据和 300 次右滑数据。

(4) 长按手势：被试在触摸板上用食指长按 50 次。被试以自己喜欢的方式以任意触点位置、触摸力度、时长和角度长按，在长按的过程中手指不可以移动，长按的时长只要求能与单次点击区分开来。被试需要保证在每次长按之前先将手指移动至目标触摸位置上方。实验共收集到 600 个长按数据片段。

(5) 拖拽手势：被试在触摸板上用食指执行拖拽手势五十次。被试在执行拖拽手势时，想象触摸板上有一个手机应用图标，并将该图标左右来回拖动。实验共收集到 300 次向左拖拽数据和 300 次向右拖拽数据。

在每两段实验之间，被试休息一分钟的时间，以避免疲劳。数据收集实验的总时长为 15 分钟。在数据收集实验后，实验者还组织了一项非正式的实验，名为“触摸踏空实验”。如图2.6所示，被试用食指点击一个白色盒子，实验者用高速双目摄像头捕捉被试点击盒子的过程。在被试连续点击数次之后，实验者要求被试闭上双眼并继续点击。数秒后实验者迅速移开盒子，此时被试的手指触摸运动会“踏空”，这一踏空的手指运动轨迹是触摸踏空实验采集的重点，用于验证触摸交互中手指的向下运动过程是否符合无约束运动模型（公式2.2）。对于每名被试，触摸踏空实验只会收集一次手指运动轨迹，这是为了避免被试知道实验意图之后，其触摸心理发生改变。



如左图所示，被试重复点击一个白色盒子。点击数次之后，如右图所示，实验者在被试进行下一次点击时迅速抽走盒子，高速摄像头记录这一“踏空”的触摸运动过程。

图 2.6 触摸踏空实验

2.3.1.2 实验设备

如图2.5所示是本实验的实验设备，包含一个 Sensei 压敏触摸板^[79]、RealSense 双目摄像头^[80]和一个 GY-91 惯性传感器。Sensei 压敏触摸板是本实验的交互表面，被试在触摸板上执行触摸交互。压敏触摸板左右宽度为 240 毫米，上下宽度为 138 毫米，包含 185×105 个传感元件，间距为 1.25 毫米，每个传感单元可以感应到大约 30000 个级别的压力，范围从 5 克到 5 千克不等。本实验通过压敏触摸板收集触摸的触点位置和压力随着时间变化的信号。

RealSense 双目摄像头可通过自带的计算机视觉方法跟踪裸手的手指位置，但是精度不够。为了提高传感精度，实验者改装了双目摄像头，将其改装成原理类似 Optitrack^[81]的针对红外反光标记点的高精度坐标跟踪系统。实验者在两个红外摄像头之间的位置上加装了一个 980 纳米波长的红外光灯泡，并用 980 纳米波长的窄通玻璃片盖住两个红外摄像头。这样一来，双目摄像头就只能看见反射红外光的物体，大大提升了图像质量。在实验中，被试在食指指甲上贴上正方形红外反射贴纸，双目摄像头可以追踪红外反射贴纸中心点的位置——红外图像中贴纸处像素为最亮的白色，通过简单的阈值方法就可以求出中心点的二维坐标，再根据双目成像原理求出手指的三维坐标，精度 σ_x 可达 0.2 毫米左右。在代码实现方面，实验者调用了 RealSense 的高帧率模式，帧速率为 300 赫兹。

商用的 GY-91 惯性传感器体积较大，为了将穿戴传感器对用户的影响降到最低，实验者自行设计电路图，将 GY-91 的核心元件 MPU8250 嵌入到一块宽度 8 毫米、厚度小于 2 毫米的线路板中。实验中，实验者用魔术贴将传感元件直接固定在被试的手指上，传感元件通过细长的杜邦线连接到 Arduino (Uno R3) 开发板上，可收集六轴的运动信号，包括三轴原始加速度和三轴角速度。其中，原始的加速

度信号是线性加速度和重力混杂的结果，系统通过 Madgwick 滤波器^[82]将原始加速度分解为线性加速度和重力方向。惯性传感单元的工作频率为 1000 赫兹，精度 σ_a 为 $0.5m/s^2$ 左右。

2.3.1.3 实验数据统计

本实验共采集了 $12 \times 5 \times 50 = 3000$ 次触摸的实验数据。实验者通过一个交互程序，人工剔除了明显错误的数据，例如，有的时候双目视觉追踪手指位置的算法会失败，导致手指位置信号发生跳变；有的时候由于被试的指甲太长，触摸板没有及时报告触摸事件。在剔除了错误数据之后，实验共记录了超过 2800 份有效的触摸交互数据。

表 2.2 本实验中触摸交互数据的统计值，括号中的数值为标准差。时间单位为毫秒，距离单位为厘米，力的单位为克，速度单位为米每秒。

统计值	单次点击	连续点击	长按手势	滑动手势	拖拽手势	总体平均
方程参数 x_0	0.84(0.34)	1.33(0.35)	0.73(0.30)	0.97(0.34)	0.92(0.33)	0.96(0.38)
方程参数 x_1	-3.89(1.50)	-3.60(1.74)	-3.19(1.55)	-4.01(1.47)	-2.90(1.72)	-3.52(1.66)
方程参数 t_1	253(60)	246(58)	263(58)	271(53)	247(64)	256(61)
触摸时间 t_c	69(11)	79(11)	71(13)	76(12)	77(13)	75(13)
抬起时间 t_{up}	176(28)	178(29)	-	206(51)	-	187(41)
触摸时长 T	107(28)	99(28)	-	130(49)	-	112(40)
触摸速度 $v(t_c)$	0.27(0.12)	0.31(0.08)	0.19(0.09)	0.26(0.10)	0.22(0.09)	0.25(0.10)
力度峰值 F	111(116)	106(58)	296(314)	101(92)	290(298)	181(229)

如表格2.2所示是触摸交互数据的简单统计。方程参数 x_0 指的是人产生触摸意图时手指与交互表面的距离，其均值为 1 厘米，即在触摸交互当中，人的非交互手指悬在交互表面上的距离大约为一厘米。触摸时间 t_c 指的是从人产生触摸意图到手指接触交互表面所需的时间，均值为 75 毫秒。触摸时长 $T(T = T_{up} - T_c)$ 指的是人的手指接触交互表面的时长，表格中只统计了短促的触摸交互的时长，均值为 112 毫秒。触摸速度 $v(t_c)$ 指的是手指接触交互表面瞬间的速度，均值为 0.25 米每秒。力度峰值 F 指的是手指接触到交互表面以后，手指向下滑压的力度的最大值，其均值为 181 克。统计上述数据对触摸交互技术有帮助，举例来说，在上一节——触摸运动的计算方法中，使用最优化方法拟合触摸运动方程时需要给各未知参数设置初始值和取值范围，初始值设定为表格2.2中的均值是较为合理的，而取值范围也可以根据上述实验结果来设计。

从表格2.2中还能看出，不同触摸交互任务下手指的运动轨迹也存在统计性差异。例如，若将单次点击、连续点击和滑动等归为短时触摸交互，将长按和拖住归为长时触摸交互，则这两类触摸的统计数据会有较大差异：总体来说，短时触摸交互的触摸速度 v_c 更大，假象目标点深度 x_1 更深，这一现象的原因可能是短时触摸交互更急促。接下来的三小节将利用本实验收集的数据，讨论触摸运动模型中无约束运动过程的合理性，评测触摸运动模型的拟合精度，并与先前模型进行比较。

2.3.2 无约束运动模型的合理性

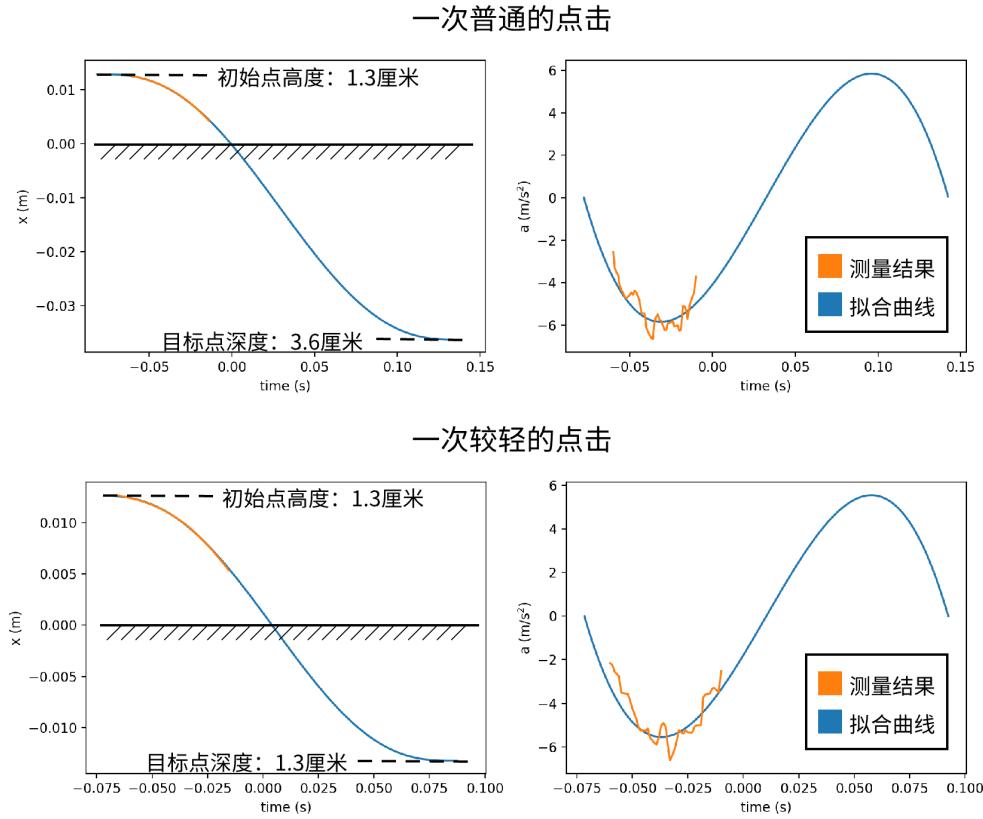
本章提到，无约束运动模型指的是：“若在触摸瞬间交互表面凭空消失，人会在规定时间 t_1 内，最平稳地将手指从初始点 x_0 移动到目标点 x_1 ”。即人在组织一次触摸运动时，其心理并非将手指带到交互表面上，而是将手指带到交互表面之下的一一个虚构点上。无约束运动模型是触摸运动模型推导过程中重要的一环，其合理性需要通过实验来求证。

在本实验的最后，实验者就追加了触摸踏空实验，收集了“触摸瞬间交互表面凭空消失”时手指的运动轨迹。由于触摸踏空实验需要对被试隐瞒实验目的，每名被试只能贡献一份实验数据，数据量较少，因此此处只作非正式的定性分析：结果显示，所有被试的手指都会运动到交互表面之下，深度在 0.43 厘米到 3.97 厘米之间，平均值为 2.26 厘米，标准差为 1.16。该结果与章节2.3.1.3中拟合的 x_1 相比更浅一些，这可能是因为真实的手指向下滑动过程受到手指结构的拉扯，不是理想的无约束运动过程。尽管如此，实验结果已经表明，被试在组织一次触摸运动时，其手指运动轨迹会经过交互表面之下的一一个虚构点 x_1 。

手指经过 x_1 后发生的事情因被试而异。在 12 名被试当中，有 7 名被试的手指会停在 x_1 上，这与无约束运动模型相符；而另外的 5 名被试的手指会在经过最低点 x_1 后折返向上，也就是说，在被试的手指达到最低点 x_1 后，手指有一个向上的加速度 ($a > 0$)，这就是章节2.1.3中更复杂的触摸运动方程的情形。前文已经提到，在这种情况下尽管触摸运动模型只是对实际情况的简化，但在计算上仍然是合理、高效的。

2.3.3 触摸运动模型的拟合精度

如图2.7所示是使用触摸运动计算模型拟合实际测量数据的两个实例，上方的两幅子图展示了一次普通力度点击的拟合效果，下方的两幅子图展示了一次较轻点击的拟合效果。左侧子图是对触摸运动位移的拟合，右侧子图是对加速度的拟合。从图中可以看出，下方触摸运动轨迹的目标点深度更浅，因此对应一次力度更轻的点击。



上方两幅子图展示了一次普通点击的位移和加速度时空曲线，下方两幅子图展示了一次较轻点击的相关曲线。从图中可以看出，拟合曲线与测量结果相符。

图 2.7 触摸运动模型的拟合实例

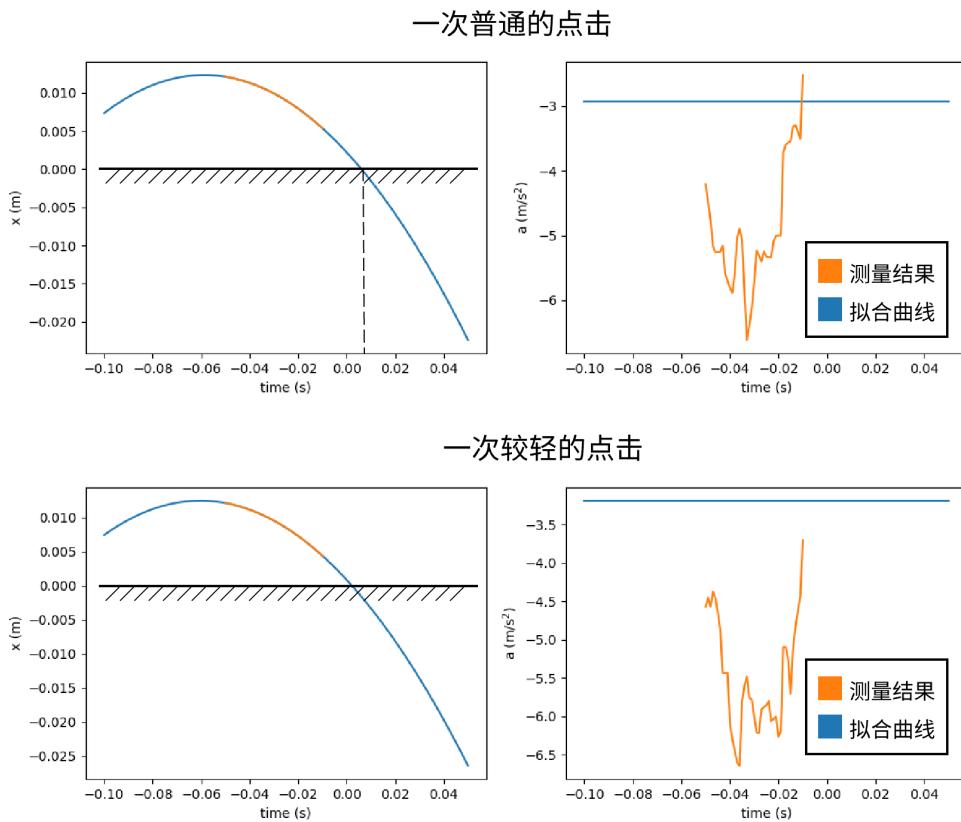
本小节通过均方根误差（RMSE）来评估触摸运动方程的拟合精度：

$$RMSE(x) = \sqrt{\frac{1}{N_x} \sum_{k=1}^n (x_m(k) - x(t_s + kT_x))^2} \quad (2.22)$$

$$RMSE(a) = \sqrt{\frac{1}{N_a} \sum_{k=1}^m (a_m(k) - a(t_s + kT_a))^2} \quad (2.23)$$

其中， $RMSE(x)$ 是触摸运动位移方程的均方根误差， $RMSE(a)$ 是加速度方程的均方根误差。在本实验的数据集中， $RMSE(x) = 8.6 \times 10^{-5} m = 0.086 mm$ ，是位移传感器标准误差 σ_x 的 43.01%，可见触摸运动模型对触摸前 50 毫秒内位移信号的拟合精度非常高。 $RMSE(a) = 0.64 m/s^2$ ，是惯性传感器标准误差 σ_a 的 127.08%。

为了对比触摸运动模型与现有模型，本小节复现了相关工作中^[56]利用抛物线拟合触摸运动的方法，该工作利用手指的空中运动轨迹预测触摸事件发生的时间，以抵消系统延迟带来的影响。如图2.8所示，抛物线能较好地拟合位移信号，其中 $RMSE(x) = 3.5 \times 10^{-5} m = 0.035 mm$ ，是位移传感器标准误差 σ_x 的 17.53%。但是，



图中展示了抛物线对两次触摸实例的拟合效果。抛物线只能较好地拟合位移信号，而无法很好地拟合加速度信号，因为触摸运动过程并非匀加速运动。

图 2.8 抛物线对触摸运动的拟合实例

抛物线无法很好地拟合加速度信号，其中 $RMSE(a) = 2.98m/s^2$ ，是惯性传感器标准差的 597.22%。这是因为，触摸运动过程并不是匀加速运动，相关工作中利用抛物线拟合触摸运动的方法是不符合原理的。相比之下，本文所提出的触摸运动模型能更好地预测触摸交互中的各项参数，例如触摸时间 t_c 、触摸速度 $v(t_c)$ 等，能为触摸交互优化技术带来更大的启发。

2.4 触摸运动模型的指导意义

触摸运动模型对触摸交互技术的优化具有指导意义，本文在两种典型的触摸交互任务下评测模型对技术的优化效果：

- **触摸指点任务：**触摸指点任务是人通过触摸交互在图形用户界面中点中目标的任务，是触摸图形用户界面中最基础的交互任务。
- **文本输入任务：**触摸文本输入任务是人通过连续触摸点击进行打字的任务，是最复杂、最快速的触摸交互任务。

如表格2.3所示，本文的第3、4、5章将介绍三项具体的触摸交互优化技术，以

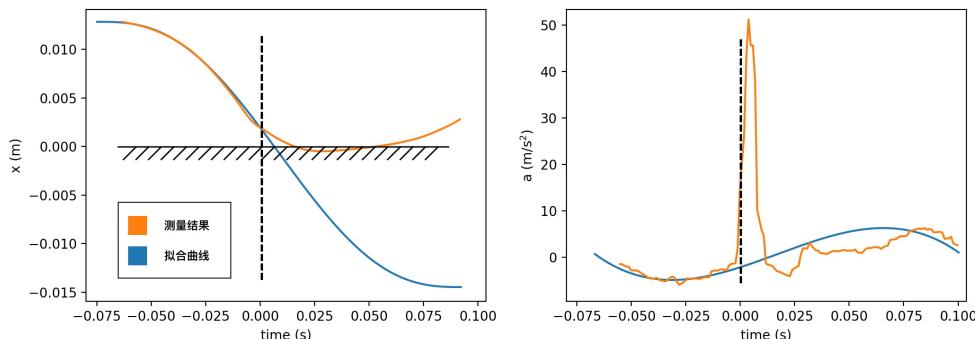
表 2.3 触摸运动模型与第3、4、5章所述触摸交互技术之间的联系。

	触摸指点任务	文本输入任务
触摸检测	第3章, 指环上的高准确低延迟触摸检测技术	
触摸手势识别		第4章, 指环上的触摸手势识别和打字技术
触摸意图推理		第5章, 平板电脑上的意图推理键盘

验证触摸运动模型在不同触摸交互技术和交互任务下的表现。本节将从理论的角度出发, 讨论触摸运动模型对触摸检测、触摸手势识别和触摸意图推理技术的指导意义, 而第3、4、5章将详细介绍技术的实施方案。

2.4.1 触摸检测

触摸检测是触摸交互中检测手指接触交互表面瞬间的技术, 是检测触摸交互是否发生的技术。触摸检测的响应性是重要改进方向, 即触摸检测的准确率应大于99%, 延迟应低于20毫秒, 以满足人对触摸交互的感官需求^[1]。



左图为一次触摸的位移时空轨迹, 右图为加速度轨迹, 黑色虚线是手指接触交互表面的瞬间, 橙线是传感器的测量结果, 蓝线是用手指接触之前的信号拟合的触摸运动方程。

图 2.9 模型为触摸检测提供判断依据

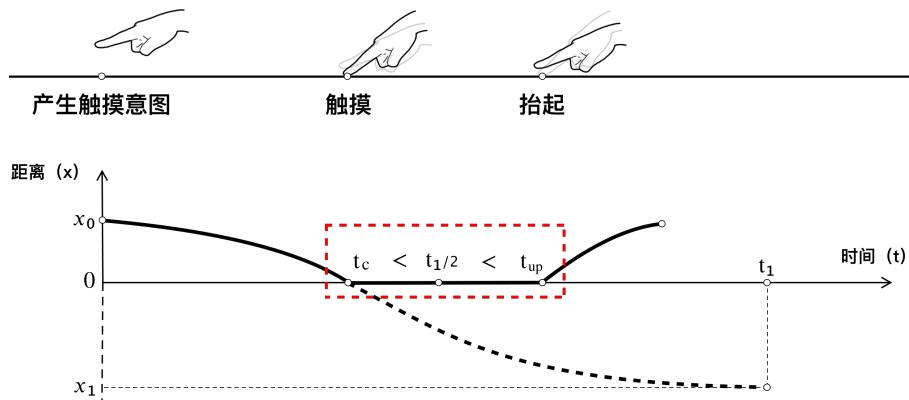
触摸运动模型为触摸检测技术提供了完整的判断依据。如图2.9所示, 触摸运动模型揭示, 触摸由手指的向下运动过程和手指瞬停构成。在手指接触交互表面之前, 手指处于向下运动过程, 其轨迹符合触摸运动方程(公式2.3); 在手指接触交互表面的瞬间, 手指由于受到交互表面的阻碍而瞬停, 使得手指位移和加速度的测量结果都大幅偏离触摸运动方程的预测结果。以位移信号为例, 当传感器测量结果与触摸运动方程相差超过三倍标准差时, 可以判定触摸已经发生:

$$|x_m(\tau_{now}) - x(\tau_{now})| > 3\sigma_x \quad (2.24)$$

其中, τ_{now} 是当前时刻, $x(\tau_{now})$ 是模型预测的手指位置, 而 $x_m(\tau_{now})$ 是传感器测量结果, σ_x 是将信号误差和拟合误差考虑在内的标准误差。从式子可以看出, 上述判据的理论准确率为三倍标准差所对应的 99.7%。更重要的是, 在保障检测准确率的前提下, 利用上述式子检测触摸的延迟是理论最低的, 因为在这一时刻之前, 运动信号所携带的信息不足以区分“触摸已发生”和“误差导致的结果”。因此, 触摸运动模型也揭示了, 触摸检测的准确性和延迟优化并非独立的优化问题, 而是此消彼长的关系。触摸运动模型可以在两者的动态平衡中找到最有益于交互体验的解, 本文将在章节3.5.1中结合人的触摸行为数据, 详细介绍触摸检测技术。

2.4.2 触摸手势识别

触摸手势识别是识别并分类触摸交互手势的技术, 常见的触摸手势有点击、长按、滑动、拖拽等等。触摸手势识别的难点是高准确、低延迟地检测触摸和抬起, 其中, **抬起检测**指的是手指接触交互表面后, 检测手指离开交互表面瞬间的技术。在准确检测触摸和抬起以后, 其它触摸手势的识别往往能通过简单的阈值方法解决, 例如, 长按手势等价于触摸和抬起之间的时间间隔超过一定阈值; 左滑手势等价于从触摸到抬起的这段时间内手指向左移动超过一定阈值。



触摸运动中点 $\frac{t_1}{2}$ 是触摸运动方程时长的一半, 大多数情况下, 触摸发生在中点之前, 抬起发生在中点之后。

图 2.10 模型为抬起检测提供辅助的判断依据

触摸运动模型为抬起检测技术提供了辅助的判断依据。如图2.10所示, 虽然触摸运动模型描述的是触摸前后极短时间内的手指运动规律, 并不涉及抬起时的规律, 但由于触摸和抬起之间存在相关性, 触摸运动模型为抬起检测提供了两点辅助的判断依据: (1) 触摸运动中点 $\frac{t_1}{2}$ 是触摸运动方程时长的一半, 大多数情况下, 触

摸发生在中点之前，抬起发生在中点之后；(2) 触摸速度与 $v(t_c)$ 与抬起速度 $v(t_{up})$ 之间存在正相关性 ($R^2=0.48$)。运用上述两条规律能够提高抬起检测的准确率，从而为触摸手势识别技术奠定基础，本文将在章节4.4.2中详细介绍触摸和抬起之间的相关性，并介绍触摸手势识别技术。

2.4.3 触摸意图推理

触摸意图推理是判断触摸是否表达交互意图的技术，也称防误触技术。自然触摸交互中人的有意触摸和无意触碰混杂，防误触技术变得重要，同时具有很大的挑战性。本文认为，最佳的防误触技术应从用户意图的层面出发，过滤任何不表达交互意图的触摸。因此，触摸意图的推理是值得研究的问题，而本文所介绍的触摸运动模型为触摸意图推理提供了判断依据。

2.4.3.1 基于触摸运动模型的触摸意图推理

由于有意触摸是用户根据其交互意图刻意组织的，而误触是人在无意识，或者是其意图与当前交互任务无关时造成的，因此有意触摸和误触的触摸运动轨迹应当有所不同。这一规律在触摸运动模型上的体现是，触摸运动方程中的众多参数（如 $x_0, x_1, t_1, v(t_c), a(\frac{t_1}{2})$ ），在有意触摸和误触上的分布是不同的：举例来说，通常情况下有意触摸的速度 $v(t_c)$ 更大、意图触摸时长 t_1 更短、手指向下运动时有明显的加速过程（加速度最值 $a(\frac{t_1}{2})$ 更大）等等。据此可基于贝叶斯推理，以较大概率区分有意触摸和误触：

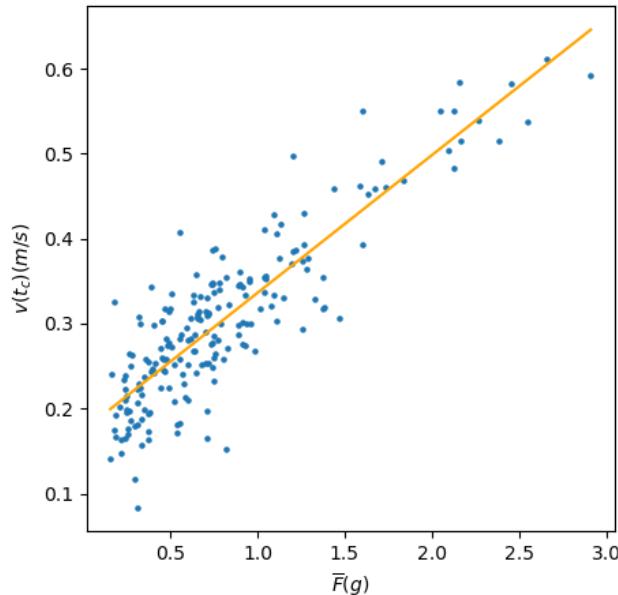
$$\begin{cases} P(I) \propto P(t_1|I)P(v(t_c)|I)P(a(\frac{t_1}{2})|I) \dots \\ P(\neg I) \propto P(t_1|\neg I)P(v(t_c)|\neg I)P(t_1|\neg I)P(a(\frac{t_1}{2})|\neg I) \dots \end{cases} \quad (2.25)$$

其中， I 表示触摸是有意的，当 $P(I) > P(\neg I)$ 时可判定触摸为有意触摸；当 $P(I) < P(\neg I)$ 时可判定触摸为误触。在理想情况下， I 或 $\neg I$ 条件概率下触摸运动方程各参数的分布可以通过大量的实验估计得；而对于实际的防误触技术而言，上述式子揭示的现实意义是，触摸运动方程中诸多的参数都可作为触摸意图的判断依据。

2.4.3.2 触摸过程与触摸力度之间的关系

为了验证本文模型为触摸意图推理提供的判据是否实际有效，本章将基于上述思路，实现一种压敏触摸屏平板电脑上的防误触键盘。压敏触摸板所采集的数据是触摸发生以后的触摸面积、触摸压力等数据，而不是触摸运动模型中所涉及

的手指位移、速度和加速度信号。在此背景下，若要对触摸运动模型加以利用，就需要探索触摸面积、压力信号与手指运动信号之间的关系，通过压敏触摸板数据间接反映触摸运动方程的参数。



触摸瞬间手指的运动速度和触摸后手指对交互平面施加的平均压力之间存在正相关的关系。

图 2.11 触摸速度与力度之间的关系

根据第2章收集的实验数据，触摸瞬间手指的速度 $v(t_c)$ 与触摸后手指对交互平面施加的平均压力 \bar{F} 之间存在关系：

$$\bar{F} = 0.16 \times v(t_c) + 0.17 \quad (2.26)$$

其中，线性拟合的决定系数 (R^2) 为 0.76，也就是说因变量 \bar{F} 的变异中，有 76% 的部分可由自变量 $v(t_c)$ 解释。包括平均压力 \bar{F} 在内，研究者发现，压力信号中还有不少特征值可以反映触摸运动方程的参数。例如，触摸瞬间手指速度 $v(t_c)$ 与压力信号的最大值 $\max(F)$ 、平均值 \bar{F} 、标准差 σ_F 均具有中等程度的相关性 ($R^2 = 0.65, 0.76, 0.67$)；人产生触摸意图时手指高度 x_0 与 $\max(F)$ 、 \bar{F} 、 σ_F 均具有弱相关性 ($R^2 = 0.43, 0.58, 0.35$)；人在组织触摸运动时计划的运动距离 $|x_1 - x_0|$ 与 \bar{F} 具有弱相关性 ($R = 0.47$)；人在组织触摸运动时计划的运动时长 t_1 与手指接触交互表面的总时长 $t_{up} - t_c$ 之间存在弱相关性 ($R^2 = 0.38$)。上述规律表明，在利用压力信号推理触摸意图的算法中，应该充分发掘压力信号的时空特征，以更好地表征触摸运动状态，从而达到推理触摸意图的目的。

2.4.3.3 面向平板电脑键盘的意图推理

平板电脑键盘上的十指打字是最快捷、最复杂的触摸交互任务之一，解决该任务的误触问题对交互效率和体验具有重要意义^[61]。在最理想的防误触能力下，平板电脑应允许用户在打字过程中将非交互手指休息在触摸屏上，而不引发误触，这样一来能大幅缓解用户在长时间打字时面临的疲劳问题。本章后续介绍的实验发现，若用户被允许将手指休息在触摸屏上，多指休息行为导致的无意触碰将占到误触总量的 75% 以上，给防误触技术带来很大的挑战。

触摸运动模型能够有效应对上述挑战，有效过滤多指休息行为导致的误触：人在将多根手指休息在触摸屏上时，多根手指的向下运动过程是协同的，因此它们的触摸运动方程参数较为类似。根据这一现象，可利用各手指触摸运动方程的时空相关性来识别多指休息行为，从而将多指休息行为导致的触摸事件判定为误触。

2.5 本章小结

本章提出了基于最优控制理论的触摸运动模型。最优控制理论与神经系统的学科交叉^[58]指出，人的肢体运动倾向于最小化其急动度的平方积分，在此假设下，本章推导出触摸运动模型的数学表达，即触摸运动过程是一个分段的五次函数（公式2.3）。进一步地，本章提出了触摸运动模型的计算方法，即使用卡尔曼滤波和最小二乘法来拟合触摸运动方程参数。最后，本章通过实验证明了触摸运动模型的合理性，也表明了触摸运动方程是对触摸运动过程的良好拟合。在本文接下来的三章中，将介绍触摸运动模型在三种典型的触摸交互技术中的应用。

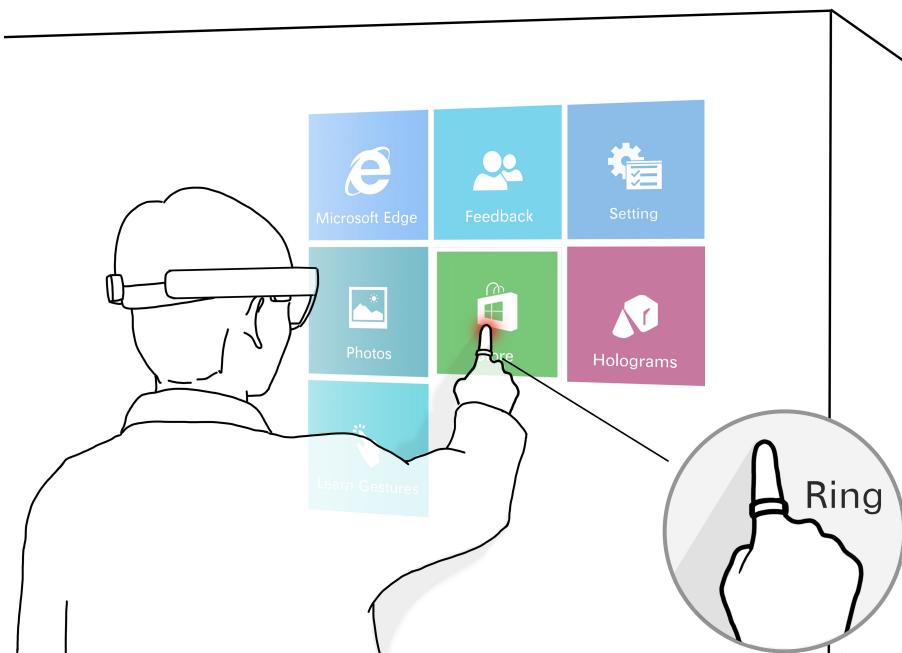
第3章 指环上的高准确低延迟触摸检测技术

3.1 引言

头戴式混合现实设备（MR 头盔，如微软的 Hololens^[4]）为新一代人机交互范式带来了丰富的可能性。MR 头盔通过前置深度摄像头感知物理环境和用户动作，通过带有显示器护目镜将三维虚拟元素渲染在物理实体之上，原则上使得 MR 头盔用户可以随时随地与数字世界交互。MR 头盔上一种也有前景、有价值的应用场景是在无源表面（如桌面、墙面）上投射虚拟用户界面，并允许人通过触摸与用户界面进行交互。上述应用场景将触摸交互——当前用户规模最大的人机交互方式，从有源触摸屏的束缚中解放出来，拓展到任意物理表面上。与目前 MR 头盔上流行的空中手势交互相比，触摸交互提供真实的触觉反馈，这是自然人机交互体验的重要组成部分。

要在 MR 头盔情境下检测无源表面的触摸交互，一个直观的想法是利用 MR 头盔的前置深度摄像头，通过视觉方法检测触摸事件。然而，视觉方法在检测触摸交互时存在固有的缺陷：首先，由于摄像头在手背的方向，摄像头在捕捉手指接触交互表面时面临严重的遮挡问题，在大多数情况下，摄像头都无法直接观察到手指接触交互表面的位置；第二，目前主流的深度摄像头是基于双目视觉的，这一类型的深度摄像头在识别交互表面的深度时精度不高，特别是当交互表面是纯色的、无明显视觉关键点的时候，误差可能达到厘米级别，对触摸交互技术的响应性带来很大挑战；第三，视觉方法的计算过程复杂，在计算资源有限的机器上运算可能会带来用户难以接受的延迟。例如，在利用视觉方法检测无源表面触摸交互的代表性工作 MRTouch^[11] 中，研究者利用 Hololens 混合现实头盔的前置摄像头检测触摸事件，其未识别率高达 3.5%，误触率高达 19.0%，端到端延迟高达 180 毫秒。这说明，基于视觉方法的无源表面触摸交互技术仍需改进。

如图3.1所示，我们设想在未来结合 MR 头盔和智能指环来使能无源表面上流畅的触摸交互体验：MR 头盔的前置摄像头主要负责识别手指的 2D 位置和姿态，而戴在手指上的惯性传感指环主要负责低延迟地检测手指是否接触到了交互表面。由于惯性传感器所提供的数据带宽较低，处理惯性传感器数据的计算通常是高效的，这确保了触摸检测的高响应性。在相关文献中，先前的工作已经探索过使用手指上佩戴的惯性传感器来增强触摸交互技术^[2,31,46]。使用智能指环感知触摸并不是本文的创新点，然而，先前工作仅对惯性传感器信号使用阈值方法来检测触摸事件，其响应性非常低下（准确率不超过 89.8%，延迟不低于 50 毫秒），因此基



结合 MR 头盔和智能指环有望在无源表面上支持流畅的触摸交互体验。

图 3.1 MR 场景下触摸交互无处不在的设想

于惯性传感指环的触摸交互技术还有很大的改进空间。

为提高指环触摸交互技术的可用性，本文提出了基于惯性传感指环的低延迟触摸检测技术。低延迟是该技术与先前工作的主要区别，该技术的检测延迟低至 10 毫秒，而人无法在触摸交互中察觉到 10 毫秒的延迟，因此该技术在响应性上提供了最佳的用户体验。本章将从各个维度介绍基于惯性传感指环的低延迟触摸检测技术：

首先，本章开展了两项用户实验来调研触摸交互指环的设计问题，分别是用户对不同触摸姿态的主观喜好程度调研（实验一 A），和用户对不同指环佩戴位置的主观喜好程度调研（实验一 B），这两项实验的结果将指导技术的交互设计。然后，本章通过用户实验采集用户佩戴智能指环进行触摸交互的数据（实验二），再利用简单的机器学习方法实现了初版的触摸检测技术，用于对比指环佩戴在不同位置上时触摸交互技术的性能。最后，本章将触摸运动模型应用在指环触摸交互技术的改进上，介绍了基于惯性传感指环的低延迟触摸检测技术。评测实验（实验三）的结果显示，该技术的触摸检测准确率高达 98.61%（精确度 98.61%，召回率 98.62%），检测延迟低至 10 毫秒。这一结果说明，基于惯性传感指环的低延迟触摸检测技术具有高响应性，同时，该技术因支持无源表面触摸交互而具有高普适性，是一项有着广阔应用前景的新技术。

3.2 相关工作

本文引言已经提到，触摸交互技术主要分为有源触摸屏技术和无源表面触摸技术两大类。有源触摸屏技术是目前最常用的触摸交互技术^[36-37,40]，但由于它被束缚在有源表面上的内禀缺陷，触摸屏技术始终不具备适应未来穿戴计算交互模态的普适性。在MR头盔等新型交互设备兴起的背景之下，无源表面触摸交互技术是有前景的研究方向。无源表面触摸交互主要有种技术实现方案，基于摄像头的视觉方法和基于震动的方法，这两种方法在本质上分别对应触摸运动的位移信号和加速度信号。

3.2.1 基于视觉的无源表面触摸交互

摄像头可以通过视觉方法观察手指触碰交互表面的时间和位置，从而检测触摸事件，支持触摸交互。当摄像头部署在场景空间中，或作为可穿戴设备部署在人身上^[10-11]，而非直接部署在交互表面上时，即可认为摄像头支持无源表面上的触摸交互，具有较高的普适性。在相关工作中，已有研究者利用激光雷达^[12]、RGB摄像头^[13-16]、红外摄像头^[17]、热成像摄像头^[18]和深度摄像头^[10-11,13,19-21]实现基于视觉方法的无源表面触摸交互技术。然而，先前基于视觉的触摸交互技术的基本原理都是直接观察手指是否接触交互表面，这种方法面临严重的遮挡问题，影响了触摸交互技术的响应性，即使在最新工作中^[11]，该方案的误触率也高达19.0%，端到端延迟高达180毫秒。本文的触摸运动模型为基于视觉的无源表面触摸交互技术带来了新思路，不同于先前技术中摄像头“直接”观察手指是否触摸表面的思路，触摸运动模型揭示：可以通过手指触摸表面时的瞬停现象“间接”地判断触摸事件，从而有效规避遮挡问题，提高触摸检测的响应性。

3.2.2 基于震动的无源表面触摸交互

在触摸交互中，当手指接触到交互表面上时，会引发手指的震动，并且发出声音。震动和声音可以被惯性传感器和麦克风所捕捉，因此，不少工作致力于探索基于震动的无源表面触摸交互，利用部署在人的手指^[5,28-29]或手腕^[30]上的传感器感知触摸。然而，先前工作仅通过阈值方法感知触摸^[2,31-32]，即当手指震动幅度或声音响度超过特定阈值时判定触摸发生，其准确率不超过89.8%，端到端延迟不低于50毫秒，不满足触摸交互的高响应性需求。本文的触摸运动模型为基于震动的无源表面触摸交互技术提供计算理论基础，模型指出，除去手指接触交互表面瞬间引发的震动以外，手指向下过程中的众多运动信号规律都有助于提高触摸检测技术的响应性。在触摸运动模型的基础之上，本文作者在2019年提出基于惯性传

感指环的低延迟触摸检测技术^[5]，亦即本章所介绍的内容，该技术将无源表面触摸交互技术的响应性提升到用户无法察觉到延迟的标准，后来有研究者跟随并改进了这份工作。例如，本文作者和 Shi 等人在同一时期提出惯性传感指环不仅可用用户检测单次触摸事件，还可以识别长按、滑动拖拽等触摸手势^[28-29]；Meier 等人提出通过两个惯性传感器，可以将设备从手指上转移到手腕上^[30]，而不会严重影响触摸交互技术的检测性能。

3.2.3 触摸姿态

触摸姿态指的是触摸交互中，人的手指接触交互表面时的全手型姿态，例如，食指指腹点击、中指指尖敲击、食指第二关节叩击、三指触摸都是不同的触摸姿态。触摸交互技术若能区分不同的触摸姿态，将能丰富触摸交互的表达力^[83-84]，然而因为目前主流的触摸屏技术无法区分不同的触摸姿态，所以相关研究一直停留在实验室阶段。随着触摸交互技术的发展，触摸姿态的多样化将被相关技术所支撑，这也是本章将就用户对不同触摸姿态的主观喜好程度进行实验调研的动机。除了触摸姿态以外，触摸压力^[85]、速度^[43-44]、切向力^[86]和手指朝向^[87-88]等额外的输入信道也能丰富触摸交互的表达力，可作为未来的研究方向。

3.3 指环触摸交互的自然性研究

由于指环上无源表面触摸交互技术对普通用户而言是陌生的，本章重视该交互方式的自然性研究。为此，本章开展了两项用户实验来探索指环触摸交互的自然性问题，分别调研了用户对不同触摸姿态和不同指环佩戴位置的主观喜好程度，从而指导本智能指环触摸检测技术的交互设计。

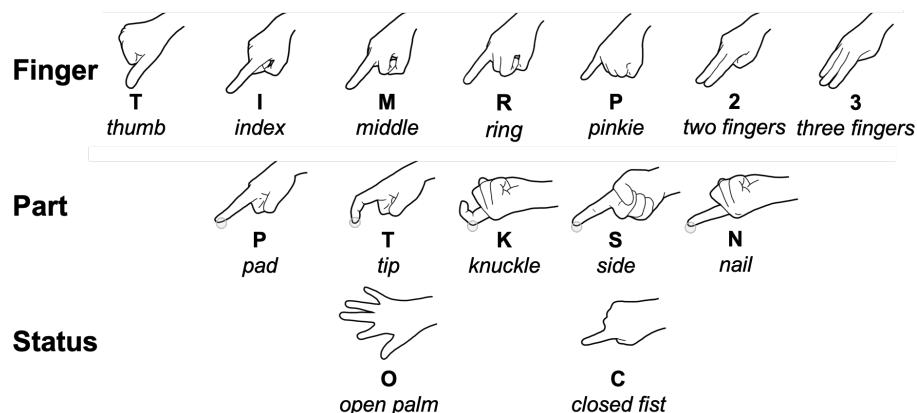
3.3.1 实验一 A：触摸姿态主观喜好程度调研

本着人机交互以人为中心的原则，好的触摸交互技术应支持用户喜好的触摸姿态，而非易于识别的触摸姿态。为此，实验一 A 通过实验和问卷调研被试对不同触摸姿态的喜好程度，以寻找触摸交互技术应支持的触摸姿态集合。首先，我们定义了一套触摸姿态全集；然后，对于触摸姿态全集中的每个姿态，我们要求被试分别将水平的桌面和垂直的墙面视为交互表面，使用该姿态进行“触摸”，亲身感受后在问卷中对该触摸姿态进行评分。最后，实验者根据被试的评分选择出最受用户喜好的触摸姿态集合。

3.3.1.1 触摸姿态全集

如图3.2所示，我们通过一个三维的设计空间来探索触摸姿态的全集，这三个维度分别是：

- 哪些手指触摸交互表面？选项包括拇指、食指、中指、无名指、小指、双指和三指。
- 手指的哪个部分触摸交互表面？参考相关工作 TapSense^[84]，选项包括指腹、指尖、手指第二关节、指侧和指甲。
- 手部状态？选项包括五指张开和握拳。



触摸姿态的设计空间有三个维度，分别是哪些手指触摸、手指的哪个部分触摸和手部状态，三个维度的笛卡尔积就是触摸姿态全集。其中，粗体字母是根据每个状态的英文描述起的缩写。

图 3.2 触摸姿态的三维设计空间

以上设计空间中三个维度的笛卡尔积即为触摸姿态的全集，共包含 $7 \times 5 \times 2 = 70$ 种触摸姿态。实验者为他们定义了缩写，例如，五指张开状态下食指指腹触摸的缩写为 IPO，其中 I 对于食指，P 对应指腹，O 对应五指张开的状态。

3.3.1.2 实验设计和过程

实验者从校园中招募了 20 名被试，其中 7 名为女性，被试的年龄是 18 岁到 27 岁不等，平均年龄为 22.0 岁。实验在两种交互表面上开展，分别是水平的桌面和垂直的墙面，它们是日常生活中场景的物体表面。实验采用组内实验设计，实验者采用拉丁方平衡被试在两种不同交互表面上的实验顺序。由于触摸姿态全集包含 70 中触摸姿态，而实验在两种交互表面上开展，每位被试在 $70 \times 2 = 140$ 种设置下体验触摸交互，并给出评分。实验问卷采用了七级李克特量表，在三个方面上收集用户对每种触摸姿态的评分：

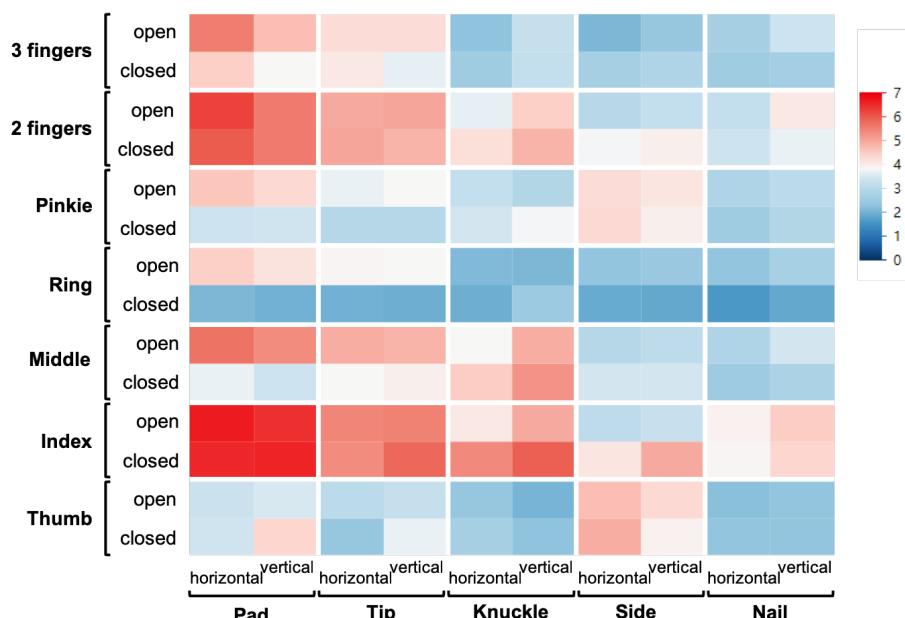
- 舒适度：以该姿态触摸的物理和心理轻松程度（1 - 不轻松，7 - 非常轻松）。
- 易记度：记住该触摸姿态的难以程度（1 - 不容易，7 - 非常容易）。

- **喜好度:** 对该触摸姿态的总体喜好程度（1 - 完全不喜欢，7 - 非常喜欢）。

在实验的最后，实验者针对以下问题对被试进行了简短的采访：(1) 您认为在本实验的触摸姿态全集之外是否有其它可用的触摸姿态？(2) 在日常使用中，您愿意记住多少种不同的触摸姿态？

实验中，被试坐在可调节的椅子上，在开始实验前，被试需要调整座椅，让自己以最舒适的姿态进行接下来的实验。对于每一个触摸姿态，实验者都会首先进行示范，然后被试亲自执行该触摸姿态三遍，然后在问卷中对该触摸姿态的舒适度、易记度和喜好度评分。由于李克特量表的主要通过相对数值对比不同的被测量值，每一次触摸之后，被试都可以通过对比来修改之前作出的评分。被试每执行十个触摸姿态，就需要休息五分钟的时间，以避免疲劳。整个实验耗时一个小时。

3.3.1.3 实验结果

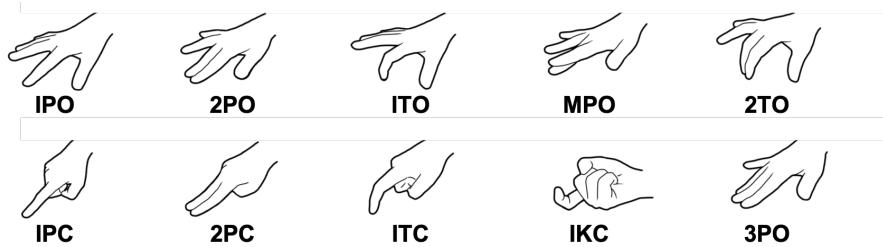


20名用户对两种交互表面上、70种不同触摸姿态的平均喜好程度，其中，1分表示完全不喜欢，7分表示非常喜欢。

图 3.3 触摸姿态主观评分热力图

图3.3展示了被试在水平和垂直交互表面上对所有70种触摸姿态的喜好程度评分，图3.4展示了被试们最喜好的触摸姿态前十名。Friedman检验显示交互表面的水平和垂直与否并不显著影响用户对各触摸姿态的喜好。在实验后的采访中，没有任何被试报告在触摸姿态全集以外存在其它可用的触摸姿态，因此，实验者认为图3.4中的就是人们最喜好的十种触摸姿态。

喜好程度排名前五的触摸姿态分别是 IPO (食指指腹触摸-张手)、IPC (食指指腹触摸-握拳)、2PO (双指指腹触摸-张手)、2PC (双指指腹触摸-握拳) 和 ITO



在人们最喜好的十种触摸姿态中，有七种姿态来自常用的单指或多指指腹点击，另有 ITO（食指指尖敲击-张手）、ITC（食指指尖敲击-握拳）和 IKC（食指第二关节叩击-握拳）。

图 3.4 人们最喜好的触摸姿态前十名

（食指指尖敲击-张手），这些姿态都是目前常用的触摸手势。喜好程度排六到十名的触摸姿态是 ITC、MPO、IKC、2TO 和 3PO，其中 MPO（中指指腹触摸-张手）和 IKC（食指第二关节叩击-握拳）值得关注：若触摸交互技术有能力区分食指指腹触摸和中指指腹触摸，部分用户将愿意让中指触摸表达特定的交互语义；食指第二关节叩击是触摸屏上不常用的触摸姿态，但近年来被华为手机用于触发截屏功能，从本实验结果来看，这一交互方式应会广受用户好评。

Friedman 检验显示，“哪些手指触摸交互表面” ($\chi^2 = 767.70, p < .0001$) 和“手指的哪个部分触摸交互表面” ($\chi^2 = 423.86, p < .0001$) 都对被试的主观评分有显著性影响。被试更喜欢用食指、中指、两指和三指进行触摸交互，不喜欢用拇指、无名指和小指触摸。被试只接受指腹点击、指尖敲击和手指第二关节叩击，厌恶手指侧面点击和手指甲敲击。在实验后的采访中，被试报告说他们平均愿意记住 7.45 (SD=2.61) 个触摸姿态，这是因为，如果交互系统中表达特定交互意图的手势太多，会给用户带来很大的认知负担。因此，实验者认为这十种最受欢迎的触摸姿态是值得进一步研究的，也值得本章的低延迟触摸检测技术去检测，而这十种触摸姿态以外的情形就不需要更多的关注了。

3.3.2 实验一 B：指环佩戴位置主观喜好程度调研

本小节介绍一个简短的用户实验，实验目的是调研用户对指环佩戴位置的主观喜好程度，其中，指环是用于触摸交互的智能指环，而非用于装饰的戒指。如图3.5(a)所示是实验所评测的九种不同的指环佩戴位置，分别是食指、中指和无名指的第一、二、三指骨。实验中，被试需要将一个嵌入了惯性传感器的指环佩戴在手指的不同位置上，在桌面上以自己喜欢的姿态触摸若干次，然后通过问卷对其体验进行主观评分。问卷采用七级李克特量表（1分 - 最差，7分 - 最好），在三个方面上收集用户对指环佩戴位置的主观看法：

- 舒适度：在此位置上佩戴指环进行触摸交互的舒服程度。

- **接受度:** 在此位置上佩戴指环的社会接受程度,比如,是否会吸引他人不必要的注意。
- **喜好度:** 对该指环佩戴位置的总体喜好程度。

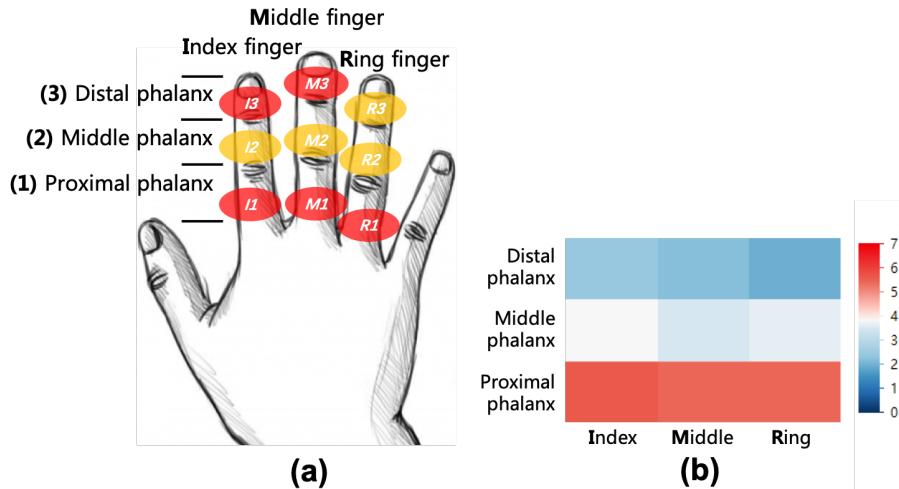


图 a 展示了实验所评测的九种不同的指环佩戴位置, 实验者为每个位置定义了缩写, 例如, I1 表示食指的第一指骨; 图 b 是被试对不同指环佩戴位置的喜好程度 (1 - 完全不喜欢, 7 - 非常喜欢)

图 3.5 指环佩戴位置主观评分热力图

实验一 A 的 20 名被试参与了本实验, 如图3.5(b) 所示, 被试更喜欢在食指第一指骨 (5.65 分)、中指第一指骨 (5.45 分) 和无名指第一指骨 (5.45 分) 上佩戴指环, 在这些指根的位置上佩戴指环是舒适、且可接受的。将指环佩戴在手指的第二指骨上的主观评分是中肯的, 而将指环佩戴在手指末端令人厌恶。实验结果表明, 人们更愿意将智能指环佩戴在手指的第一指骨上, 但具体是戴在食指、中指还是无名指上都差异不大。因此, 基于惯性传感指环的触摸检测技术应尽可能通过指根上的运动传感信号感知触摸事件。

3.4 基于机器学习的触摸检测技术

3.4.1 实验二: 收集触摸数据

实验二采集了被试进行触摸交互的惯性传感器数据和头戴式摄像头数据, 实验的动机是为两项后续工作提供数据支持: 一是基于摄像头的触摸姿态识别, 二是基于惯性传感信号的触摸检测技术。

3.4.1.1 实验设计和过程

实验者从校园中招募了 12 名学生作为被试，其中有 4 名女性，被试的年龄从 20 岁到 29 岁不等，平均年龄为 23.1 岁。实验分为两大部分，每名被试分别在水平的桌面和垂直的墙面上进行实验。实验采用了组内实验设计，实验者采用拉丁方平衡了水平或垂直表面的实验顺序。

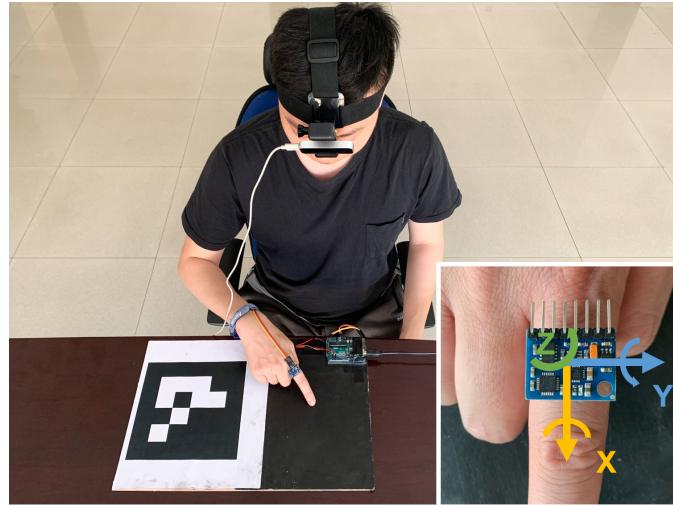
对于每种交互表面，需要进行五段实验：如图3.5手指上的红色区域所示，每段实验中被试分别在五个不同的位置上佩戴智能指环，这五个位置分别是 I1（食指第一指骨）、M1（中指第一指骨）、R1（无名指第一指骨）、I3（食指第三指骨）和 M3（中指第三指骨）。选择 I1、M1、R1 等手指第一指骨位置是因为实验一的结果表明，用户更喜好将智能指环佩戴在这些位置上；而加入 I3 和 M3 这两个位置是因为实验者猜测，手指末端上惯性传感器所采样的触摸信号的震动强度更大，利用该信号感知触摸事件的能力可能会更强，实验者需要通过实验对比来验证此猜测是否成立。实验舍弃了采集指环佩戴在 I2、M2、R2 和 R3 位置上的触摸数据，这是因为冗余的、时间过长的实验可能会导致被试的疲劳，影响数据的可靠性。

每段实验包含十组，在每组实验中，被试分别以如图3.4所示的十种触摸姿态接触交互表面，各 20 次。实验者要求被试尽可能用自然的方式进行触摸交互，每名被试总共进行了 $2 \times 5 \times 10 \times 20 = 2000$ 次触摸。最后，实验者还收集了一系列空中手势作为负样本，被试将惯性传感指环佩戴在不同的位置，执行画圈、画方框、空中点击等手势。采集空中手势的过程中，被试不允许让手指接触到其它物体。对于每个指环佩戴位置，需要采集持续一分钟的空中手势数据。在实验中，被试要在每 200 次触摸后休息两分钟，以避免疲劳问题。整个实验持续了一个半小时。

3.4.1.2 实验设备

如图3.6所示是实验二的实验设置。被试在手指上佩戴了惯性传感指环，头上戴着一种商用手型追踪摄像头（LeapMotion），被试在一块自制的超低延迟触摸板上进行触摸交互。实验通过指环中的惯性传感单元采集加速度和角速度信号，通过 LeapMotion 采集手部的骨骼点数据，通过触摸板采集接触与否的真值。

惯性传感指环由一个九轴的惯性传感单元（GY-85）和常规的戒指粘合而成，实验者一共制作了四个大小不一的惯性传感指环，以适应不同被试的手指大小，惯性传感单元通过杜邦线连接在 Arduino（Uno R3）开发板上。实验者用魔术贴将杜邦线固定在被试的手腕上，以求在最大程度上减轻杜邦线对用户触摸交互造成的影响。超低延迟触摸板是一块覆盖有导电墨水的木板，当手指触摸到木板时，木板与用户的手指形成耦合电容，电容会增加，实验者利用这一现象来判断手指接



实验二的实验设置，被试佩戴惯性传感指环接触触摸屏。子图展示了惯性传感指环及其信号所在的坐标系。

图 3.6 实验设置

触木板的精确时间^[89]。

触摸板也通过杜邦线与 Arduino 连接，它和惯性传感指环共享相同的时间戳，他们的频率都为 200 赫兹。高速摄像头的图像分析表明，从人触摸木板到 Arduino 开发版上的信号灯亮起所需时间不超过 5 毫秒，也就是触摸板的延迟在 5 毫秒以下。触摸板旁边是一个二维码标识符，用于准确标记交互表面所在位置，以配合 LeapMotion 计算人的手指上每个关节与交互表面之间的距离。LeapMotion 的传感频率为 60 赫兹，其与 Arduino 之间的传输延迟大概为 20 毫秒。LeapMotion 在工作原理上依赖其自带的红外照明，一般而言，它时而会出现手型估计失败的情况，为此实验者非常注重实验的关照条件：实验在灯光明亮且避免阳光的环境中进行，触摸板上的导电墨水也专门采用吸收红外光的墨水，为摄像头提供了黑色背景，以上实验环境使得 LeapMotion 的传感精度得到保障。

3.4.1.3 实验数据统计

本实验共采集了 $12 \times 2000 = 24000$ 次触摸的实验数据。实验者开发了一个交互式程序，用以人工剔除明显错误的数据，例如，有时候指甲太长的用户接触触摸板时，触摸板无法及时报告触摸事件。在剔除错误数据之后，实验共产生了超过 23900 份有效的数据作为触摸的正样本，实验者从空中手势中随机抽取负样本，正样本和负样本的数量相同。

3.4.2 基于机器学习的触摸姿态分类

本小节将介绍基于支持向量机（SVM）的触摸姿态分类方法，该方法较为直观，其目的只是评测采用视觉方法分类触摸姿态的可行性，而非本文的贡献。若该方法的分类准确率达到可观的水平，说明本章提出的 MR 头盔加智能指环的组合不仅可以检测触摸事件，还能区分不同的触摸姿态。

3.4.2.1 技术方案

本小节参考了相关论文^[90]提取手部骨骼特征，特征包括手指尖与交互平面的距离、手指尖之间的距离、手指尖与手掌的距离和手指尖与手掌平面之间的夹角，这些值拼接成 19 维的手型特征向量，实验者根据此特征向量提取方法训练 SVM 模型，该模型即可分类不同的触摸姿态。

3.4.2.2 评测结果

如表3.1所示，实验者采用留一（被试）交叉验证法评估触摸姿态分类的准确率。IPO、IPC、2PC、IKC 这四种最常用的触摸姿态的四分类准确率达到 99.0%，这说明触摸姿态四分类具有较高的可行性；在四分类的基础上，外加 2PO、MPO 和 3PO 这三种触摸姿态，得到七分类的准确率为 88.5%，这说明触摸姿态七分类的准确率仍在可接受范围内；然而，所有常用触摸姿态的十分类准确率却很不理想，准确率仅为 77.0%，这说明就目前的视觉技术而言，要准确分类十个或者以上的触摸姿态仍比较困难。

表 3.1 触摸姿态分类的准确率（括号中数值为标准差）

	无	四分类	七分类	十分类
水平桌面	99.1% (1.3%)	89.5% (3.9%)	76.4% (6.8%)	
垂直墙面	99.0% (1.4%)	87.6% (4.8%)	77.6% (6.7%)	

上述实验结果表明，头戴式前置摄像头大约可以稳健地区分四到七种不同的触摸姿态。随着手部跟踪技术的发展^[91-93]，作者认为，通过触摸姿态的多样化来丰富触摸交互将很快成为可能。

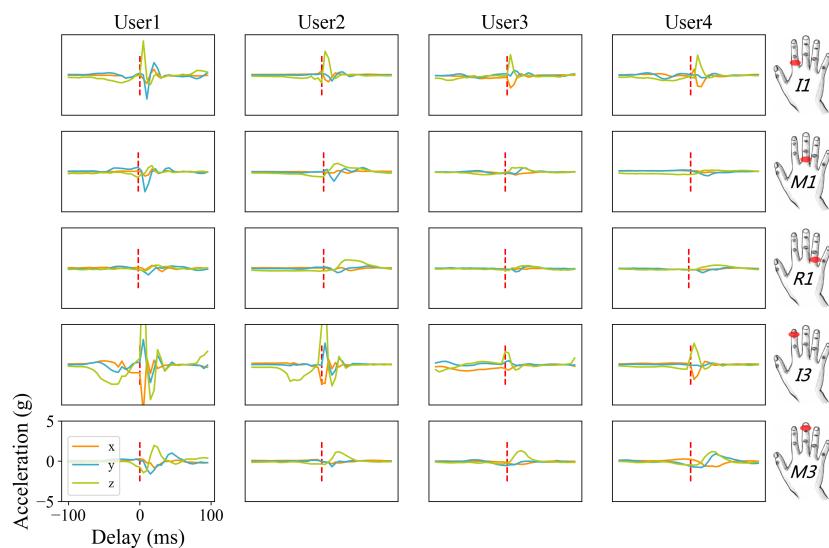
3.4.3 基于机器学习的触摸检测技术

本小节将介绍基于支持向量机（SVM）的触摸检测技术。支持向量机是一种简单的机器学习方法，具有开发敏捷的特点，因此，本小节的目的并不是实现终版触摸检测技术，而是通过简单的机器学习方法探索指环触摸检测技术中尚待确定

的因素。

3.4.3.1 数据分析

惯性传感指环的中能收集到的数据包括指环的加速度、角速度和磁力计数据。由于磁力计数据的刷新频率只有 8 赫兹，实验者舍弃了磁力计数据。原始的加速度数据中混杂了重力，实验者通过 Madgwick 过滤器^[82]将原始加速度信号分解为加速度和重力。最后，经过处理的实验数据共包含九个维度，包括三轴加速度、三轴角速度和三轴重力。

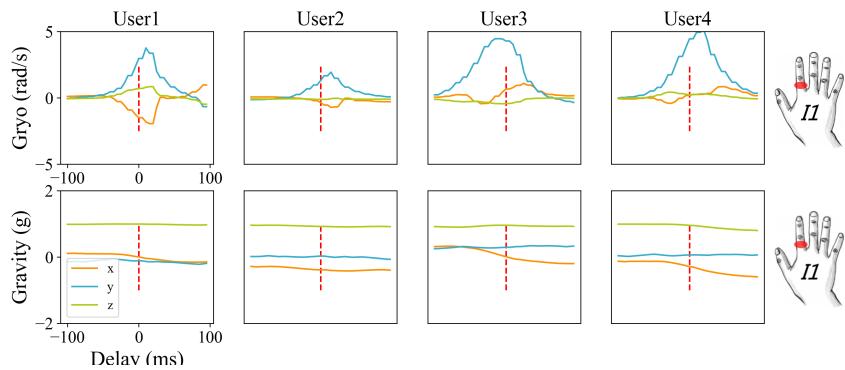


图中展示了被试执行食指指腹触摸时的惯性传感器加速度信号，四纵列子图对应四个不同的被试，五横列子图对应了五种不同的指环佩戴位置。

图 3.7 加速度信号图示

如图3.7所示，本小节以食指指腹点击这一触摸姿态为例分析数据规律。从图中可以看出，触摸导致的震动会传导到各个手指的各个位置上，形成一个幅度较大、较为尖锐的波峰。当指环佩戴在引发触摸的手指上时（如食指第一指骨、食指第三指骨），震动带来的波峰会在触摸后 10 毫秒内达到峰值；当指环佩戴在其它手指上时，震动带来的波峰则会在 20 毫米内达到峰值。对于特定的指环佩戴位置，不同被试数据中的加速度波形是类似的。通过观察波形图，实验者推断，触摸发生前后加速度的最大值、最小值、平均值、偏度和峰度等多种特征都可能有助于对触摸的感知。

如图3.8所示是角速度和重力的波形。在触摸发生前后，角速度和重力变化也是有规律的，例如，由于触摸交互时人的手掌大致平行于交互表面，所以所有被试的重力数据都是相似的，若有一段手部运动的重力数据波形与图中所示的形态不



图中展示了被试执行食指指腹触摸时惯性传感器的角速度和重力信号，其中惯性传感指环佩戴在食指第一指骨上，四纵列子图对应四个不同的被试。

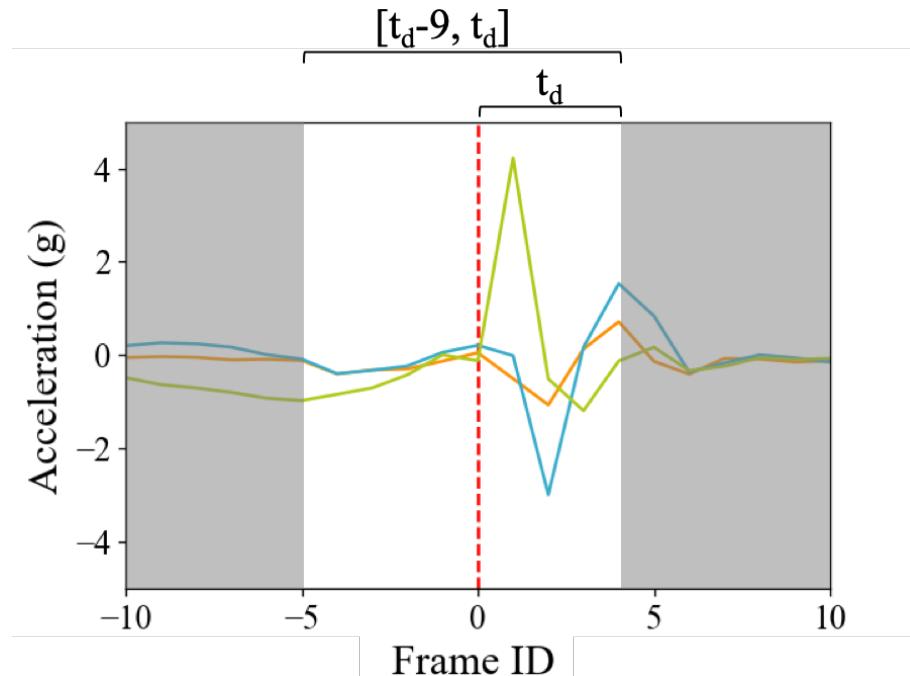
图 3.8 角速度和重力信号图示

符，说明它更有可能属于空中手势交互，而不是触摸交互。上述观察表明，来自惯性传感指环的可用信息是非常丰富的，若利用机器学习方法综合考虑各方面的有效信息，应能在触摸检测的性能上超越先前工作中的阈值方法^[2,31-32]。

3.4.3.2 触摸检测分类器

触摸检测分类器是根据一段时间内惯性传感信号，判断此刻是否发生触摸事件的二分类模型，是基于机器学习的触摸检测技术的核心。分类器从十帧（50毫秒）的传感信号中提取特征，对于九轴惯性传感器数据中的每一个维度，分类器计算其最大值、最小值、均值、偏度和峰度，然后将这些值拼接成 45 维的数组，作为机器学习模型的特征向量。

如何选择信号采样的时间窗口，是一个值得探讨的问题。触摸导致的震动传输到惯性传感器的位置需要一定的时间，如果在震动抵达惯性传感器之前就贸然判定触摸事件是否发生，准确率会受到很大的影响。因此，在本小节所介绍的触摸检测技术中，检测延迟并不是越低越好，检测延迟和准确率是此消彼长的权衡惯性，实验者必须通过模拟实验来确定延迟的最佳取值。如图3.9所示，本小节将 t_d ($0 < t_d < 10$) 定义为触摸检测分类器的延迟，设触摸发生在第 0 帧，则分类器由时间窗口 $[t_d - 9, t_d]$ 的信号所提取的特征训练而得。 t_d 的值的选取是一个值得权衡的问题， t_d 的值越大，分类器就越能够利用触摸实际发生后的震动信号，分类的准确率将会提高，但与此同时检测的延迟也会增加。因此，实验者必须通过测试来给模型选定合适的延迟 t_d 。若给定 t_d ，分类器的训练过程如下：从时间窗口 $[t_d - 9, t_d]$ 的惯性传感信号中提取特征作为正样本；为了避免提前汇报触摸事件，需从时间窗口 $[-14, -5]$ 的惯性传感信号中提取特征作为负样本；此外，需要从空中手势中截取信号，提取特征作为负样本。最后，针对以上正负样本数据集运行



触摸检测分类器是判断当前时刻是否发生触摸事件的二分类模型，其原理是将十帧（50毫秒）的传感信号提取成特征向量，而后使用支持向量机训练得模型。

图 3.9 触摸检测分类器

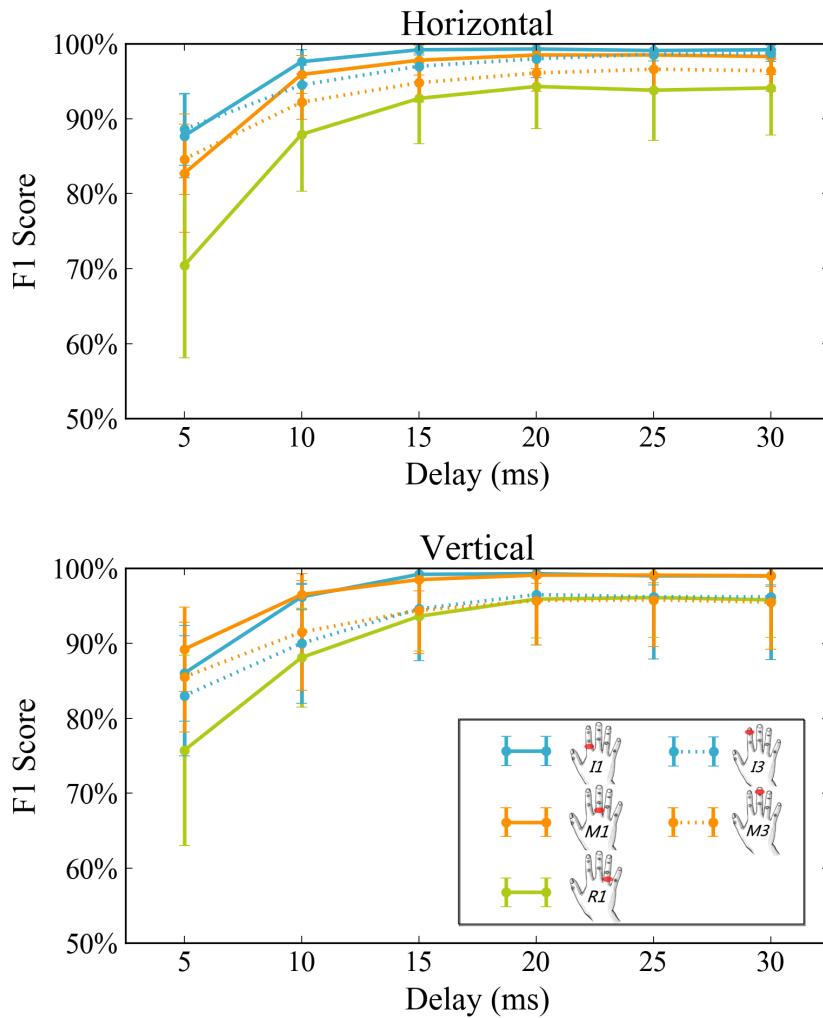
支持向量机来训练分类器模型。

3.4.3.3 检测延迟优化

上一小节提到，触摸检测分类器的检测延迟 t_d 是由技术实现者自行决定的，是一个检测准确率和延迟之间权衡的问题：一方面，检测延迟越低，触摸检测的准确率就越低；另一方面，检测延迟过高，又会直接影响用户体验。因此，如图3.10所示，实验者通过实验评测了 t_d 取不同的值时，触摸检测技术的准确率。从图中可以看出，总体上来说更高的检测延迟的确可以提高模型的预测准确率，但是当延迟提高至 20 毫秒以后，再往上延长检测时间的意义就不大了：混合方差分析表明，在 20 毫秒以内，模型的延迟 t_d 对模型预测准确率有显著性影响 ($F_{3,33} = 133.4, p < .0001$)；在 20 毫秒以后，图中的曲线就开始收敛了 ($F_{2,22} = 0.011, p = .99$)。这说明，触摸检测分类器的检测延迟 t_d 应设置为 20 毫秒，既能支持较高的检测准确率 (99.3%)，又能避免用户察觉到延迟的存在。

3.4.3.4 指环佩戴位置对比

在确定 20 毫秒的检测延迟效果最佳之后，实验者对比了 20 毫秒延迟下，指环佩戴在不同位置上的检测准确率。结果如表3.2所示，当指环佩戴在 I1（食指第一指骨）上时，触摸检测分类器的性能是最好的，准确率（F1 综合评价指标）达



在本小节所介绍的触摸检测技术中，检测准确率和延迟之间是一个此消彼长的权衡问题，图中展示了指环佩戴在不同位置上时，检测准确率和延迟之间的关系。

图 3.10 触摸检测准确率与延迟之间的关系

指环位置		I1	M1	R1	I3	M3
水平	精准率	99.7%(0.5%)	99.2%(1.0%)	97.6%(1.9%)	99.1%(1.3%)	98.3%(1.4%)
	召回率	98.9%(1.6%)	97.9%(3.0%)	91.7%(9.2%)	97.1%(4.4%)	94.1%(4.1%)
	F1 分数	99.3%(1.0%)	98.5%(1.8%)	94.3%(5.6%)	98.0%(2.5%)	96.1%(2.5%)
垂直	精准率	99.7%(0.6%)	99.3%(1.1%)	98.1%(1.5%)	98.3%(2.1%)	98.4%(1.9%)
	召回率	99.0%(0.9%)	98.8%(1.8%)	94.0%(8.4%)	95.4%(10.8%)	93.7%(9.8%)
	F1 分数	99.3%(0.6%)	99.1%(1.1%)	95.9%(5.3%)	96.5%(6.7%)	95.7%(5.9%)

表 3.2 指环佩戴在不同位置上时的触摸检测准确率（触摸检测延迟为 20 毫秒）

到 99.3%；次好的指环佩戴位置是 M1（中指第一指骨），准确率为 98.8%。方差分

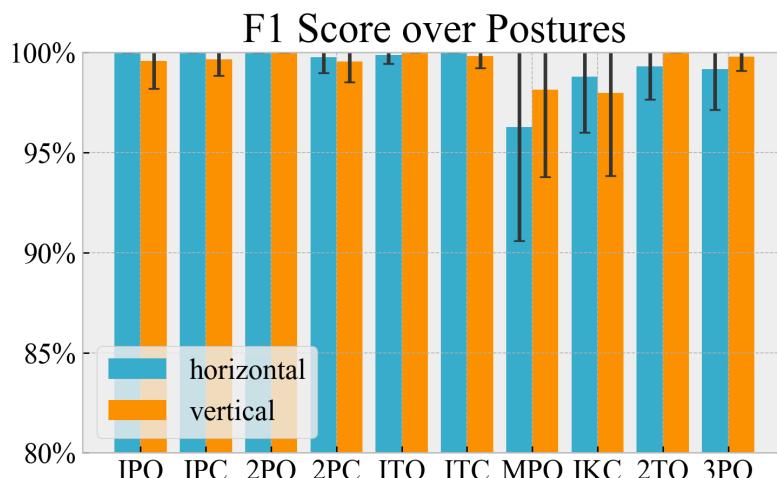
析显示，指环佩戴位置对检测准确率有显著性影响 ($F_{4,44} = 6.45, p < .001$)，而交互表面的方向则对检测准确率没有显著影响 ($F_{1,11} = 0.09, p = .76$)。后验检测显示 I1 指环佩戴位置显著优于 R1 ($p < .001$)、I3 ($p < .005$) 和 M3 ($p < .005$)；M1 显著优于 R1 ($p < .005$)、I3 ($p < .05$) 和 M3 ($p < .01$)。

实验一曾提到，实验者将指环佩戴位置 I3 和 M3（食指、中指第三指骨）也加入了实验序列，其原因是实验者猜测此处能传感到更强的触摸震动，有利于触摸检测的准确率。然而，从本实验的结果看来，此猜测被证伪了，当指环佩戴在位置 I3 和 M3 上时触摸检测的准确率反而降低了。通过观察实验数据，实验者找到了两点原因：首先，虽然位置 I3 和 M3 上的惯性传感器能传感到更强的振动，但是其噪声也被放大了；其次，一根手指触摸交互表面所产生的震动不能很好地传递到另一根手指的指尖。

综合本小节实验结果和本章中对指环佩戴位置的主观喜好程度调研，将惯性传感指环佩戴在 I1 或 M1 位置（食指或中指第一指骨）上是最佳选择，既能提高指环触摸检测技术的检测准确率，又能满足用户的主观喜好。

3.4.3.5 综合评测

以上结果表明，触摸检测分类器在 20 毫秒的检测延迟下，在用户将惯性传感指环佩戴在食指第一指骨上时，性能最佳。因此，我们在此最佳设置下对触摸检测分类器做全面的评测。



图中展示了触摸检测分类器对十种常用触摸姿态的触摸事件检测准确率，其中误差线表示标准差。

图 3.11 分类器对各触摸姿态的检测准确率

图3.11展示了不同触摸姿态下触摸检测的准确率（F1 综合评价指标）。当被试执行敲击姿态 MPO（中指指腹点击）和 IKC（食指第二关节叩击）时，准确率超

过 95%，其它触摸姿态的触摸事件检测准确率均接近 100%。这一结果说明，佩戴在食指上的惯性传感指环在感知食指触摸产生的震动时十分灵敏，而在感知其它手指触摸产生的震动时灵敏度有所下降，但其准确率仍然是可以接受的。

实验者设计了两种基线技术（baseline）用于与本小节所介绍的触摸检测技术做对比，这两种技术分别是基于震动和阈值的方法，和基于视觉的方法：（1）实验者仿照先前工作^[2,31]实现基于震动和阈值的触摸检测技术，并通过对本章所用数据集进行模拟实验来找到技术中所涉及参数的最优阈值；（2）实验者仿照先前工作^[11,19,21-22]实现了基于视觉方法的触摸检测技术，当摄像头检测到手指指尖与交互表面的距离下降到 10 毫米以下时，即汇报触摸事件。图3.12现实了本小节所介绍方法与两种基线技术之间的对比，方差分析表面本方法显著提高了触摸检测的精准率 ($F_{1,11} = 10.4, p < .001$) 和召回率 ($F_{1,11} = 59.8, p < .0001$, $F_{1,11} = 124.7, p < .0001$)。

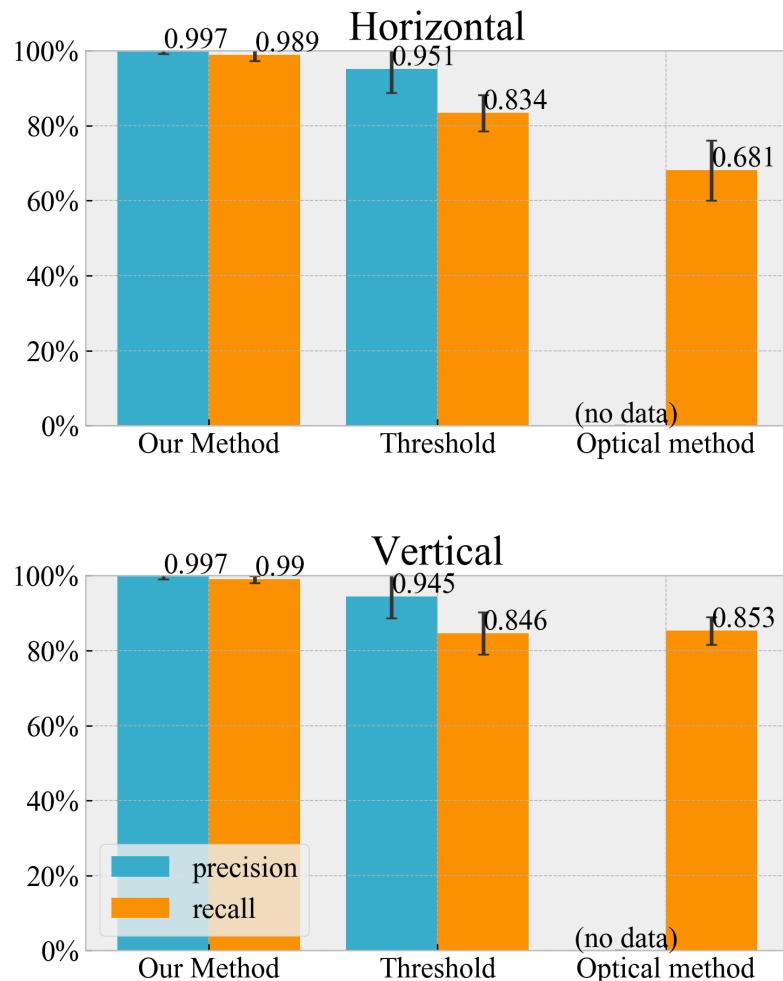
3.4.3.6 触摸检测算法流程

触摸检测分类器只能就一段惯性传感信号，判断当前是否发生了触摸事件，然而，该分类器还不足以实时感知触摸，这是因为：（1）若时刻通过分类器判断触摸事件，分类器将在手指触摸交互表面后重复报告多次触摸事件；（2）虽然分类器的预测准确率高达 99.3%，但是传感器的运行频率为 200 赫兹，即每秒钟会调用 200 次分类器进行判断，这样一来误触的概率还是很大的。为了解决上述问题，实验者设计了一种基于触摸检测分类器的触摸检测算法：

- 如果在过去的十帧（50 毫秒）内分类器预测过触摸事件，则算法不会重复报告触摸事件。
- 只有当分类器连续两帧判定为触摸时，算法才会报告触摸事件。

上述算法会让检测延迟多一帧，但是，它能够大大降低误触的概率，并且不会重复报告多次触摸事件。本章将在实验三中通过实际的触摸交互任务评测该触摸检测算法的性能。

本小节介绍了基于支持向量机（SVM）的触摸检测技术，该方法只是简单版本的触摸检测技术，其目的在于探索清楚基于惯性传感指环的触摸交互技术中的诸多位置问题，为基于触摸运动模型的触摸交互技术打下基础。本小节所得出的结论如下所述：首先，先前工作中采用阈值方法判断触摸事件的方式是低效的，惯性传感指环能提供非常多有用的信息，用于支持低延迟的触摸检测技术；其次，将惯性传感指环佩戴在食指或中指的第一指骨上是最佳选择，该指环佩戴位置既符合用户的主观偏好，有能保证最高的触摸检测准确率；第三，触摸导致的震动能在 10 毫秒内传递到接触交互表面的手指上，能在 20 毫秒内传递到相邻手指上。其



图中展示了本触摸检测技术在平均精准率和召回率上的对比，其中，误差线表示标准误差。实验中没有负样本可用于评测视觉方法的精准率。

图 3.12 本触摸检测技术与 baseline 的对比

中，接触交互表面的手指的震感强烈，相邻手指第一指骨上的震感也强烈，但是相邻手指第三指骨处的震动已经不足以支撑高准确率的触摸检测了。上述结论为下一节内容，基于触摸运动模型的低延迟触摸检测技术奠定基础。

3.5 基于触摸运动模型的触摸检测技术

3.5.1 基于触摸运动模型的触摸检测技术

基于触摸运动模型的触摸检测技术适用于运动传感信号，包括位移、速度和加速度信号。其中，基于摄像头的视觉方法可传感位移信号，惯性传感器可收集加速度信号。本章所提出的 MR 头盔加上惯性传感指环的智能设备系统可以获取触摸运动的位移和加速度信号，可用于支持低延迟触摸检测技术。本小节将分三种传感信道的情况讨论触摸运动模型如何指导触摸检测技术，这三种情况分别是：(1)

仅有位移传感信号；(2) 仅有加速度传感信号；(3) 融合位移和加速度信号。

(1) 若仅有位移传感信号：在追踪手指的过程中，设当前时刻的时间为 t_{now} ，用 $t \in [t_{now} - 60, t_{now} - 10]$ 毫秒时间内的位移信号拟合无约束运动方程 $x(\tau) = x_0 + (x_1 - x_0)(6\tau^5 - 15\tau^4 + 10\tau^3)$ ，使用信赖阙方法^[78]进行拟合，若拟合成功，则说明此时手指正处于无约束运动过程中，需要观察手指是否停止运动，以判断手指是否接触到了交互表面。若当前位移测量值 $x_m(\tau_{now})$ 与运动方程的预测值 $x(\tau_{now})$ 相差超过传感信号的三倍标准误差时，可判定触摸事件发生：

$$|x_m(\tau_{now}) - x(\tau_{now})| > 3\sigma_x \quad (3.1)$$

若在成功拟合无约束运动方程后，始终不发上述公式所述情况，说明人的手指做出了空中点击的动作，而非真正的触摸点击。实验表明，触摸瞬间手指速度的经典值为 $0.15m/s$ ，若位移传感信号的标准误差为 0.5 毫米，则根据上述公式，触摸事件的检测延迟为 10 毫秒左右。

(2) 若仅有加速度传感信号：在记录手指加速度的过程中，设当前时刻的时间为 t_{now} ，用 $t \in [t_{now} - 60, t_{now} - 10]$ 毫秒时间内的加速度信号拟合无约束运动的加速度方程 $a(\tau) = (x_1 - x_0)(120\tau^3 - 180\tau^2 + 60\tau)$ ，使用 SLSQP 算法进行拟合，若拟合成功，则说明此时手指正处于无约束运动过程中，需要观察手指运动是否达到一个很大的加速度，以判断手指是否接触到了交互表面。若当前加速度测量值 $a_m(\tau_{now})$ 与运动方程的预测值 $a(\tau_{now})$ 相差超过传感信号的三倍标准误差时，可判定触摸事件发生：

$$|a_m(\tau_{now}) - a(\tau_{now})| > 3\sigma_a \quad (3.2)$$

由于触摸运动的位移方程在触摸瞬间 $t = t_c$ 不可导，公式揭示此时的加速度无限大。但实际上，手指不是理想的刚体，所以加速度不可能是无限大，而仅仅是相当大。若惯性传感器紧贴在食指的第一关节骨上，它的加速度会在碰撞发生后的 10 毫秒内达峰值，触摸事件的检测延迟应为 10 毫秒左右。

(3) 融合位移和加速度传感信号：

若同时有位移信号和加速度信号，可以融合两种信号，在保持准确率不变的情况下降低触摸事件检测的延迟。在当前时刻 t_{now} ，位移信号服从正态分布 $N(x(\tau_{now}), \sigma_x)$ ，设位移信号的概率密度函数为 $f_x(x_m(\tau_{now}))$ ：

$$f_x(x_m(\tau_{now})) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left(-\frac{(x_m(\tau_{now}) - x(\tau_{now}))^2}{2\sigma_x^2}\right) \quad (3.3)$$

同理, 设加速度信号误差的概率密度函数为 $f_a(a_m(\tau_{now}))$, 则公式3.1、公式3.2分别对应:

$$\int_{|x_m(\tau_{now}) - x(\tau_{now})| < 3\sigma_x} f_x(x_m(\tau_{now})) > 99.7\% \quad (3.4)$$

$$\int_{|a_m(\tau_{now}) - a(\tau_{now})| < 3\sigma_a} f_a(a_m(\tau_{now})) > 99.7\% \quad (3.5)$$

由于卡尔曼滤波已经有效利用位移信号和加速度信号的互信息, 降低了两者的标准误差, 此时可认为位移的误差与加速度的误差相互独立, 因此下述公式成立时可判定触摸事件发生:

$$\int_{\frac{|x_m(\tau_{now}) - x(\tau_{now})|^2}{(3\sigma_x)^2} + \frac{|a_m(\tau_{now}) - a(\tau_{now})|^2}{(3\sigma_a)^2} < 1} f_x(x_m(\tau_{now})) f_a(a_m(\tau_{now})) > 99.7\% \quad (3.6)$$

简化以上公式得到, 当下述公式成立时可判定触摸事件发生:

$$\frac{|x_m(\tau_{now}) - x(\tau_{now})|^2}{(3\sigma_x)^2} + \frac{|a_m(\tau_{now}) - a(\tau_{now})|^2}{(3\sigma_a)^2} > 1 \quad (3.7)$$

以上就是触摸运动模型对触摸检测技术的计算理论指导, 根据上述数学推导, 接下来将介绍触摸检测的算法流程:

1. 通过摄像头和视觉的方法实时采集用户手指的位移信号, 通过佩戴在手指上的惯性传感指环采集手指的加速度信号, 每一时刻都截取从当前时刻往前推 100 毫秒内的信号数据, 作后续处理。
2. 根据 Madgwick 算法^[82]计算指环加速度信号在重力方向上的投影, 然后根据第二章所介绍的触摸运动模型的计算方法, 拟合当前的无约束运动方程, 若拟合成功, 则说明此时手指正处于无约束运动过程中, 跳至下一步骤。
3. 开始检测手指的位移和加速度信号是否满足公式3.7, 若满足, 则判定触摸事件已发生, 若在无约束运动方程拟合成功后 20 毫秒内, 公式3.7未得到满足, 则回到上一步骤

以上就是基于触摸运动模型的触摸检测算法, 接下来一节将通过真实的触摸

交互任务评测本技术，同时对比本小节中提到的一系列初探的或先前工作的触摸检测技术。

3.5.2 实验三：评测本节触摸检测技术

本实验通过实际的触摸交互应用程序评测本章所提出的触摸检测技术，实验共对比了三种不同的触摸检测技术，分别是：（1）基线（baseline）技术，先前工作中基于视觉的触摸检测技术^[11]；（2）本章第3.7节中介绍的，基于机器学习的触摸检测技术；（3）本章第3.8节中介绍的，基于触摸运动模型的低延迟触摸检测技术。为了让上述技术在尽可能苛刻的条件下接受评测，实验者选取了一款名为“别碰白键”的游戏作为实验的交互任务，该游戏要求玩家进行尽可能快速的连续触摸点击。

3.5.2.1 实验设计和过程

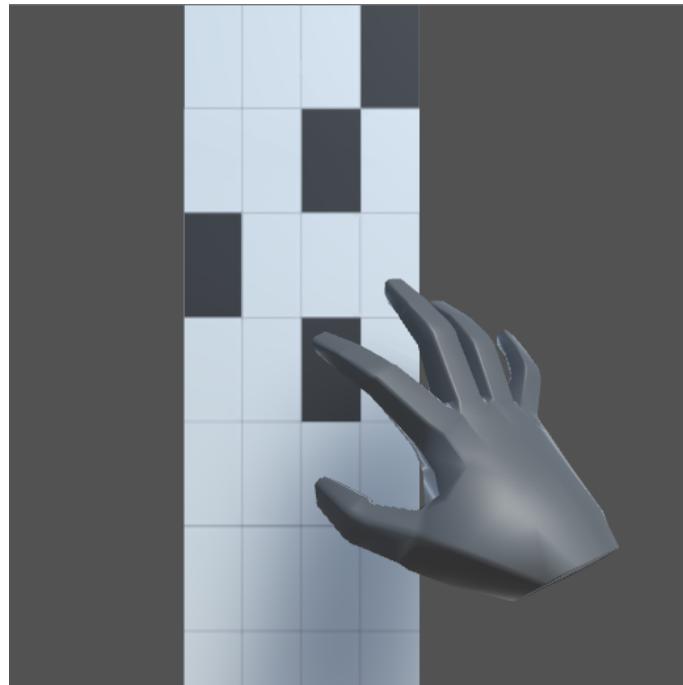
实验者从校园中招募了12名被试，其中3名为女性，被试的年龄从20岁到28岁不等，平均年龄为23.2岁。这一批被试未参与之前的任何一个用户实验。实验共分为两大部分，被试分别在水平的桌面和垂直的墙面上进行游戏。每一部分的实验又包含三段实验，每段实验采用不同的触摸检测技术。采取组内实验设计，即每一名被试都要使用不同的触摸检测技术进行实验，为了避免学习效应，实验者采用拉丁方来平衡不同触摸检测技术的实验顺序。

实验中，被试像实验二那样佩戴惯性传感指环，同时，在头上佩戴手型追踪摄像头LeapMotion，实验设置图同实验二的3.6。被试将指环佩戴在食指第一关节上，通过最常用的触摸姿态，即食指指腹点击进行触摸交互。如图3.13所示，本实验在一个显示器上展示了“别碰白键”这款游戏，同时也在游戏中渲染了被试的虚拟手，显控比为1:1。“别碰白键”这款游戏的目标是尽可能快速地通过触摸来点击从屏幕顶部出现的黑色按键，同时避开白色按键。每点中一个黑色按键，新一行的按键就会从屏幕的顶部弹出；如果被试不慎点中了白色按键，屏幕就会闪烁一秒以报告错误。游戏中，被试需要点击100个黑色按键来结束游戏，游戏的目标是尽快点击完这100个黑色按键。

本实验的总时长为20分钟，我们通过触摸检测的准确率、延迟，和用户完成任务的时间来评价三种不同触摸测试技术的性能。

3.5.2.2 实验结果

表3.3展示了实验的评测结果。在检测准确率（F1综合评价指标）方面，方差分析表明，无论是基于触摸运动模型的触摸检测技术（ $F_{1,11} = 87.9, p < .001$ ），还



图中所示为本实验的交互任务：别碰白键，这是一款要求用户以尽可能快的速度进行连续触摸的游戏，能够有效测试触摸检测技术的性能。

图 3.13 实验任务图示

	基于视觉方法 ^[11]	基于机器学习	基于触摸运动模型
精准率	85.42%(10.42%)	98.62%(2.50%)	99.32(0.74%)
召回率	84.08%(9.24%)	98.61%(1.33%)	99.17(0.88%)
任务完成时间（秒）	44.30(19.19)	35.74(13.69)	36.21(14.82)
检测延迟（毫秒）	5.48(15.07)	9.11(3.41)	9.02(3.16)

表 3.3 表格对比了不同触摸检测技术的性能，括号中的数值为标准差。

是基于机器学习的方法 ($F_{1,11} = 53.6, p < .001$)，都显著优于先前工作中基于视觉的方法。通过分析视觉方法中的错误案例能发现，本实验的交互任务要求被试连续快速触摸，基于视觉的方法不能很好的处理这种情况，例如，基于视觉的方法认为手指离开交互表面超过 15 毫米才算一次触摸动作完成，但在连续快速触摸任务中，连续两次触摸之间被试的手指可能从未离开交互表面超过 15 毫米，这影响了第二次触摸事件的检测。而在本章所提出的两种触摸检测技术的比较中，基于触摸运动模型的触摸检测技术的准确率超过 99%，显著高于基于机器学习方法的准确率 ($F_{1,11} = 5.6, p < .05$)。数据分析发现，这一差异主要来源于对较轻的触摸的检测，基于触摸运动模型的触摸检测技术检测更准确，这可能是因为基于模型的方法更好地利用了触摸之前手指向下运动的信息。

在被试完成实验任务的时间方面，方差分析表明，基于触摸运动模型的触摸检测技术 ($F_{1,11} = 7.9, p < .05$) 和基于机器学习的方法 ($F_{1,11} = 8.3, p < .05$)，都显著优于先前工作中基于视觉的方法。这是因为在视觉方法中，过高的检测错误率让被试经常需要纠正错误，延误了完成任务的时间。而在本章所提出的两种触摸检测技术的比较中，被试完成任务的时间没有显著性差异。

在检测延迟这方面，基于机器学习或触摸运动模型的触摸检测技术的延迟都稳定在 10 毫秒以内。而基于视觉的方法的检测延迟则非常不稳定，从表格中可以看到，视觉方法的检测延迟平均值为 5.48 毫秒，但是其标准差却高达 15.07 毫秒。有的时候视觉方法甚至会提前汇报触摸事件，导致负延迟的情况，这是因为视觉方法规定，当手指与交互表面下降到 10 毫米阈值以下时，即报告触摸事件。采访被试发现，没有被试能在基于机器学习或触摸运动模型的触摸检测技术下察觉到延迟的存在，但在基于视觉方法的技术中，则能明显感受到延迟忽快忽慢。

综合上述实验结果，基于触摸运动模型的触摸检测技术具有很高的检测准确率，显著高于先前工作，也显著高于基于简单机器学习的方法。而在检测延迟这方面，基于触摸运动模型的触摸检测技术也显著优于先前工作。这说明，基于惯性传感指环的触摸检测技术很具有实用前景，值得更多的关注。

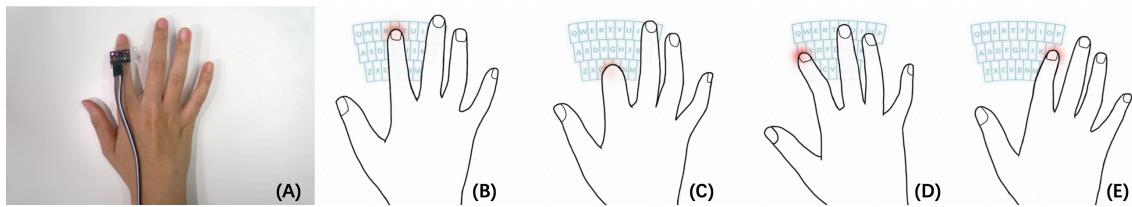
3.6 本章小结

任意无源表面上的触摸交互很可能是未来的一种主要人机交互方式，头戴式混合现实系统可以在普通物体表面上叠加交互界面，使得这些物体表面上的触摸交互成为可能。先前工作已经提出利用 MR 头盔的前置摄像头跟踪手型^[11]，但是相关工作在判断手指是否接触到交互表面上时遇到了困难。本章内容是第一份专注于优化惯性传感指环触摸检测技术的研究。评测结果表明，基于触摸运动模型的指环触摸检测技术可以在 10 毫秒以内的低延迟下检测触摸事件，召回率高达 99.17%，误触率仅为 0.68%。低延迟是本触摸交互技术的核心贡献之一，这是因为，人在触摸交互中无法察觉到低于 24 毫秒的延迟^[1]，而本技术的延迟仅为 10 毫秒，在响应性上提供了最佳的用户体验。从原理的角度来看，基于触摸运动模型的触摸检测技术性能更优有其底层逻辑：(1) 触摸运动模型充分利用了多帧信息，降低单帧信号误差带来的影响；(2) 无需直接观测触摸点，有效回避视觉遮挡问题^[11]。

第4章 指环上的触摸手势识别与打字技术

4.1 引言

头戴式显示设备（如混合现实、虚拟现实头盔）的文本输入交互亟需改进，现有的头戴式显示设备文本输入法利用头动^[94]、空中手势^[95]，或者在头盔上的触摸屏上滑动^[96-97]来输入文本。然而，基于头动和空中手势的文本输入技术会导致用户疲劳，不利于长时间文本输入；而基于头盔触摸屏的文本输入方法效率较低（每分钟打字速度小于10个英文单词）。与智能手机上广泛使用的基于触摸的软键盘相比，头戴式显示设备上的文本输入法在效率和用户体验上都较差。为了解决这一问题，本章提出了基于惯性传感指环的无源表面文本输入技术，并将其命名为**智能打字指环**：用户只需戴上一枚惯性传感指环，就能在普通的物理表面（如桌面）上快速打字，打字速度达到每分钟输入20.59个英文单词。由于智能打字指环支持无源表面上的文本输入，交互区域和显示区域解耦，该技术原则上适用于头戴式显示设备、智能电视、大屏幕等新型显示设备。



用户将惯性传感指环佩戴在食指的第二指骨上，并将手腕放在无源交互表面上（图A），他想象在他的食指可触及的范围内有一个二十六键键盘（图B-E），然后打字。智能打字指环根据惯性传感数据识别触摸事件，并预测用户想要的单词。用户从任意显示设备（如显示器）接受视觉反馈。

图4.1 智能打字指环的交互流程

如图4.1所示是智能打字指环的交互流程图，这是一种基于六轴惯性传感器指环的无源表面文本输入技术。智能打字指环适用于任意类似桌面的交互表面，交互表面要求是刚体的、平坦的、且足够宽敞，至少需要能够容纳一个手机的大小。用户通过触摸无源交互表面键入字符，并从单词推荐列表中选中所需的单词。在使用过程中，用户将惯性传感指环佩戴在食指的第二指骨上，并通过外部显示器（如混合现实头盔、增强现实头盔、智能电视）接受视觉反馈。用户将手腕轻放在交互表面上，想象其食指可触及区域内有一个二十六键键盘，然后打字。软件系统通过惯性传感指环的信号检测触摸事件，并根据指环在手指连续几次触摸中的方向预测用户所需的单词。

相比于头戴式显示设备上已有的文本输入方案，智能打字指环具有多重优势。首先，智能打字指环仅通过惯性传感指环采集用户打字信号，传感器的生产成本低、功耗低；其次，智能打字指环使用方便，该设备是可穿戴的，允许用户在任意类似桌面的无源表面上打字，具有很高的普适性；第三，用户在使用智能打字指环打字时不需要将注意力集中在手部，而可以专注在显示设备上，这一点对虚拟现实头盔和智能电视交互场景来说都是很重要的^[98]；最后，智能打字指环是易于学习的，由于它借鉴了手机软键盘的二十六键布局，保留了用户在日常打字过程中积累的关于打字的肌肉记忆。

现有文献尚未探索使用智能指环支持无源表面文本输入的可行性，这其中有一些挑战：首先，仅使用智能指环，而不辅以其它传感设备时，手指在交互表面上的 2D 位置估计往往是不准确的（误差大于 1 厘米）^[99]；其次，虽然上一章提到，可以利用惯性传感指环识别无源表面上的触摸事件，但仍有许多触摸手势的识别未被支持，包括触摸表面的手指何时抬起、长按和滑动手势等等。为了解决上述问题并构建智能指环打字技术，我们进行了三项用户实验，用于回答以下三个研究问题：

(1) RQ1：如何利用惯性传感指环识别多种触摸手势？ 实验一收集了被试的触摸交互数据，包括点击、长按和滑动等触摸手势。上一章所述的基于惯性传感指环的触摸识别技术已经支持了对点击触摸的支持，在此基础上，研究者根据触摸运动模型，支持对手指抬起事件的检测，准确率（F1 综合评价指标）达到 99.50%。进一步地，基于对触摸事件和手指抬起事件的检测，系统可以通过简单的阈值方法识别其它触摸手势，包括长按、向左滑动和向右滑动。

(2) RQ2：如何设计文本输入法的候选词解码器？ 实验二收集了被试使用智能打字指环在无源表面上打字的数据，基于数据，实验者分析了被试的打字行为，从而设计了智能打字指环的候选词解码器。候选词解码器是根据用户字符输入序列猜测用户所需单词的贝叶斯解码器。结合 trigram^[28] 语言模型，解码器在模拟实验中有 79..0% 的概率将用户所需单词预测为第一位候选词，有 94.5% 概率将用户所需单词预测为前五位候选词。结果表明，该解码器有能力支持基于选词的文本输入方法。此外，实验者通过观察实验数据发现，被试们的打字行为之间存在较大差异，这启发实验者探索个性化文本输入模型。

(3) RQ3：用户使用智能打字指环打字的速度有多快？ 实验三在使用通用候选词解码器和个性化解码器两种情况下，评测了被试使用智能打字指环进行文本输入的打字效率。实验三是为期五天的用户实验，评测了被试使用智能打字指环打字时的中短期学习效应。结果表明，无论是使用通用候选词解码器还是个性化解

码器作为，智能打字指环的表现都是良好的。平均而言，被试在 40 分钟的练习之后就能达到每分钟输入 13.74 个英文单词的打字速度，在为期五天，每天打字 40 分钟的练习过后，其打字速度达到每分钟输入 20.59 个英文单词。从第三天开始，个性化候选词解码器呈现出由于通用候选词解码器的趋势 ($p=.09$)。

本章工作的贡献有三个方面：(1) 首先，根据触摸运动模型，实现了在无源交互表面上的多手势触摸交互，包括对手指抬起事件、长按、左滑和右滑的识别，识别准确率超过 99%；(2) 其次，本章介绍了仅使用单个惯性传感指环作为输入设备，支持无源表面上 26 键文本输入的方法；(3) 本章通过实验证明，智能打字指环为混合显示头盔、智能电视等新型显示设备提供了实用、高效的文本输入方案。

4.2 相关工作

4.2.1 可穿戴文本输入

随着可穿戴设备与技术的发展，越来越多的研究者对可穿戴设备上的文本输入方法产生兴趣，利用可穿戴设备支持文本输入的用途很广泛，适用于多种新型显示设备，例如增强现实头盔和虚拟现实头盔^[94,97]、大屏幕显示器^[100-101]、智能手表^[102-103]，等等。

4.2.1.1 智能手表文本输入

利用可穿戴设备上的触摸屏是研发可穿戴文本输入的最直接方法，而智能手表是其代表性设备平台。“胖手指问题”是智能手表文本输入技术的主要挑战^[9]：由于智能手表触摸屏的空间有限，手表触摸屏中单个按键太小会导致用户无法准确点击的问题。对于字符级别的文本输入，一些技术采用多次点击确定单个字母^[102,104-105]或区别对待不同手指点击^[106]的方法来克服“胖手指问题”，还有的技术则采取规避手指遮挡的策略^[107-108]。尽管与上述工作类似的方法有很多，但是从信息论的角度来说，只要是小触摸屏上的文本输入方法，就会面临屏幕大小制约了文本输入速度上限的问题^[9]。此外，基于触摸屏的智能手表文本输入需要双手的参与，使其适用场景受限。

另有一部分智能手表上的文本输入技术不依赖触摸屏。COMPASS 用户通过旋转智能手表的表圈进行文本输入^[103]，输入速度达到每分钟 12.5 个英文单词。WrisText 用户通过手腕动作进行文本输入^[109]，输入速度达到每分钟 15.2 个英文单词。以上技术不遵循广泛使用的文本输入交互规范（例如 26 键键盘和触摸交互），因此具有较高的学习门槛，不容易学习和走向实用。

4.2.1.2 空中手势文本输入

大部分基于空中手势的文本输入技术都依赖视觉方法追踪手部运动，它们适用于可穿戴场景，例如，混合现实头盔的前置摄像头可用于追踪手部运动。手部追踪技术支持了基于一笔画^[110]、目标选择^[100]、滑动手势键盘^[111]和手写^[101]的空中手势文本输入。对比上述工作发现，Qwerty二十六键键盘是基于目标选择的空中手势文本输入中最好的布局，可使文本输入速率达到每分钟18.9个英文单词^[100]。滑动手势键盘比基于目标选择的技术更适用于空中文本输入任务，输入速率达到每分钟28.1个英文单词^[112]。然而，由于缺少物理支撑，基于空中手势的文本输入技术都会导致用户疲劳，影响长时间文本输入的体验。此外，文本输入需要高精度的输入信号，需要使用Optitrack^[81]等昂贵的摄像头设备进行准确的手部跟踪。因此，在目前的商用混合显示设备中采用空中手势文本输入方案是不切实际的。

4.2.1.3 智能手套文本输入

不少相关工作专门为文本输入设计了智能手套。基于智能手套的文本输入技术分为两种，分别是利用（1）空中手写，和（2）手指之间的敲击来输入文本。基于空中手写的技术是易于学习的^[113-114]，但是存在用户易疲劳的问题，而且，凡是基于手写的文本输入技术，都存在每分钟输入15个英文单词的打字速度上限^[115]。基于手指间敲击的文本输入法依赖特定的输入规则^[116-121]，因此学习门槛高是一个大问题。DigiTouch^[121]是一个巧妙的设计，它将二十六键键盘的每个字母拆分映射到不同手指的各个指骨上，拇指点击相应指骨来输入文本。DigiTouch用户通过三个小时的练习，能够达到每分钟输入16.0个英文单词的输入速度，其结果是同类型工作中较为出色的。

4.2.1.4 智能指环文本输入

相关工作中基于智能指环的文本输入方法较少。一部分工作提出利用智能指环识别26个英文字母^[122-124]，然而，这些工作仍专注于提升将手势识别成英文字母的准确率上，准确率不超过93%，这一准确率是无法支撑流畅的文本输入体验的。ThumbText是一款佩戴在食指上的小型触摸屏^[125]，用户使用拇指分两步点击触摸屏来输入一个字母，专家用户的打字速度为每分钟输入11.4个英文单词。TipText是一个在食指第一指骨上的虚构软键盘^[126]，键盘布局是一个2乘3的网格，布局针对盲打进行过优化，专家用户的打字速度为每分钟13.3个英文单词。RotoSwype是一种基于惯性传感指环的指向的空中滑动输入键盘^[95]，经过五天的训练，用户的平均打字速度可以达到每分钟输入14.8个英文单词。

TypingRing^[127]与本章所介绍的技术最为相似，同为基于指环的无源表面文本输入技术，用户分两个步骤打字。TypingRing 将二十六键键盘分为九个区域，每个区域包含三个按键。用户首先通过平移他们的手来选中一个区域，然后通过用食指、中指或者无名指触摸来三选一选出想要的字母。TypingRing 存在一些不足：首先，分两个步骤进行的文本输入并不直观，打字效率也很低，该技术的专家用户打字速度为每分钟输入 10.0 个英文单词；其次，论文没有公布实验细节，无法判断实验是否规范；第三，该技术的触摸识别部分准确率仅为 80%，无法支持流程的文本输入体验。

4.2.1.5 可穿戴文本输入小结

除了以上研究工作以外，还有其它输入方式可用于支持可穿戴设备的文本输入。例如，头动^[94]、倾斜设备^[128]、控制触摸压力^[129]、眼动^[130-132]和一维触摸屏^[97]都可用于打字。然而，基于触摸和二十六键布局的键盘仍然是当前智能设备上最常用的文本输入系统。因此，实用性研究应该尝试在触摸和 Qwerty 键盘布局的基础上研发文本输入技术，而本章所述的基于惯性传感指环的无源表面文本输入则是可行方法之一。

4.2.2 智能指环交互

智能指环是人机交互领域中一个快速发展的方向，2015 年一份文献综述 Digital Digits^[133]对智能指环相关工作做了全面分析，指出智能指环的三大功能。首先，智能指环的主要功能是传感超出人类手指感知范围的事物，例如手势识别^[124,134-138]。第二，智能指环可向佩戴者提供信息，例如触觉反馈^[139-140]。第三，人的手指可通过智能指环控制或发送信息到外部对象，例如智能指环控制系统^[25,141-143]。大多数研究，包括本章所介绍内容，都是将指环作为传感设备。本节相关工作介绍将从手指跟踪和触摸检测两个方面讨论智能指环的识别能力。

4.2.2.1 手型跟踪技术

声学^[144]和磁力学^[25-27]都可以作为智能指环手型追踪技术的基本原理，但只能用于检测智能指环与其它手部穿戴设备之间的相对位置，其识别精度在 4.4 毫米到 13 毫米不等。AuraRing^[27]是基于磁力学的智能指环手型跟踪技术中效果最好的，其识别精度为 4.4 毫米，然后，AuraRing 的配置不仅包含一个指环，还包含一个带有多个传感线圈的腕带，这使得它在便携性上不如单指环系统一样出色。视觉方法同样可以支持智能指环手型追踪技术^[47,145-146]，但他们存在遮挡、高功耗和高计算资源消耗等问题。惯性传感单元可以嵌入在智能指环中，用于跟踪手型，

具有成本低、体积小、功耗低等优点。MIDS 由智能手表外加两个带有惯性传感器的指环组成^[2]，LightRing 由一个用于测量手指屈伸程度的红外接近光传感器和一个用于测量手指选择的陀螺仪组成^[99]，它们都可以用户追踪手指在无源表面上的位置，但是精度降低且需要校准。在我们的工作中，我们尝试使用单个 6 轴 IMU 来估计手指的弯曲和旋转。

4.2.2.2 触摸识别技术

智能指环可支持无源表面触摸识别技术，即识别手指触摸无源表面的交互事件。先前基于指环的工作仅利用阈值方法识别触摸事件^[2,31,46]，准确率不超过 89.8%。直到如上一章所述，本文作者利于触摸运动模型改进指环触摸识别技术^[5]，指环触摸识别的准确率才达到实用的级别（99.3%）。然而，上一章内容仅支持了对手指接触无源表面瞬间的识别，不包括对手指触摸交互表面以后抬起事件的识别，也不包括对长按、滑动等多样化触摸手势的识别。与其它可穿戴文本输入技术^[9,94,126]一样，本章将要介绍的智能打字指环文本输入技术也需要包括长按和滑动在内的多种触摸手势来支持文本输入应用程序。因此，为了构建智能打字指环文本输入技术，本章还需先解决手指抬起事件的检测和其它触摸手势的识别问题。

4.2.2.3 智能指环交互小结

综合上述文献可以发现：(1) 即使通过个性化的校准^[99]，单个智能指环跟踪手指位置仍是不精确的，误差大于 1 厘米；(2) 尽管本文第二章已经攻克了基于指环的低延迟触摸识别技术，但是对更多触摸手势的识别仍然需要进一步探索。鉴于现有工作存在以上不足，在构建智能打字指环之前本章必须先解决两个先决技术：(1) 首先，使用贝叶斯解码器从不准确的输入信号中联想用户想要的单词。与许多其它依赖智能联想的文本输入技术一样^[95,109,112,147]，智能打字指环也无法支持词汇表外（OOV）的单词；(2) 第二，研发可以检测无源表面上多样化触摸手势的技术，触摸手势包括手指抬起、长按、滑动，等等。

4.2.3 贝叶斯解码器

贝叶斯方法被广泛应用于文本输入解码器，用于通过输入信号联想用户所需单词^[148]，其基本思想是根据词语出现的概率（语言模型）和触摸点的分布（触点模型）来估计词库中每个单词作为用户想要单词的概率，从而将概率最大的单词上屏，或者将概率最大的若干候选词展示出来供用户选择。对于语言模型，本章将评估一元、二元和三元模型^[149]对单词联想准确率的影响。对于触点模型，大多数文本输入研究假设每个按键上的触摸点云服从二维高斯分布，这是绝对触点模

型^[148,150]。然而，BlindType 这篇论文^[98]指出，当用户在文本输入中手眼分离时，连续触点之间的向量也服从二维高斯分布，这是相对触点模型。

4.3 智能打字指环的交互设计

本节首先介绍无源表面指环文本输入技术（智能打字指环）的交互设计，包括硬件配置、键盘布局和交互流程；然后，本节再介绍设计准则，即为什么要采取所述交互设计。在硬件上，智能打字指环是嵌入了六轴惯性传感的智能指环，交互时用户将指环佩戴在食指第二指骨上。如图4.1A 所示是智能打字指环的交互流程：用户将手腕放在桌子上，想象在其食指可触及区域内有一个 Qwerty 二十六键键盘。用户不能真的在桌面上看到一个二十六键的键盘布局，而是仅凭想象判断二十六个英文字母按键的位置，因此本节将此键盘成为想象的二十六键键盘。用户通过额外的显示设备来查看键盘布局。为了减少用户定位想象键盘上每个字母位置的认知负担，键盘的 26 个按键被设计成占满了食指的整个可触及区域，图4.1B-E 展示了想象键盘的边界：第一排键位位于食指指腹的最远触摸范围内；第三排键位位于食指指尖轻敲的位置上；键盘的左右边界刚好在用户转动手腕时食指能够轻松够到的范围内。

4.3.1 键盘布局



图中展示了智能打字指环的键盘布局，Rank-1 表示最有可能的候选词。字母 Q、A、Z、P、M 的中心坐标位置是 (0,0)、(0.2,1)、(0.8,2)、(9,0) 和 (6.8)。

图 4.2 智能打字指环的键盘布局

用户从外部显示器（例如增强现实头盔、显示器或智能电视）观看如图4.2所

示的键盘布局作为视觉反馈。键盘由二十六键布局图和候选词选择区域组成。二十六键布局图的作用是提醒不会盲打的用户 26 个字母按键所在的位置。在文本输入的过程中，系统将始终预测当前输入序列对应的前五名候选词，显示在候选词选择区域上，从左到右的候选词排序是第四名、第二名、第一名、第三名、第五名候选词。也就是说，最有可能的候选词是放在中间的，次有可能的候选词放在两边，这一设置旨在节省选择单词所需的时间。在文本输入的过程中，用户可以将注意力集中在显示器上，而不需要将注意力放在自己的手上。

4.3.2 交互流程

用户按以下步骤输入单词：

1. 在想象的二十六键键盘上按顺序点击所想单词的每个字母。
2. 长按 200 毫秒触发候选词选择功能，先不抬起手指，这时在 Rank-1 候选词上会出现一个光标。
3. (可选) 若 Rank-1 候选词不是想要的单词，通过左右滑动控制光标选择其它单词。当指环沿着竖直轴旋转 7.5° 时，光标移动一格。
4. 抬起手指选中光标所在的候选词上屏。

用户通过左滑手势删除单词。与许多其它依赖智能联系的文本输入技术一样^[94,103,112]，如果用户所需单词不在候选词列表中，用户需要删除错误的单词，重新输入。在使用智能打字指环输入文本时有一个限制，由于指环上的惯性传感器只能传感食指的指向，但不能传感手指的平移，因此用户需要将手腕放在交互表面上不要平移，只通过转动手腕来触及不同的按键。只有有了这个约束，惯性传感指环才可能区分同一行的不同键位。用户需要一些时间的熟悉来适应这一要求，本章将在后面讨论熟悉阶段的学习效应。

4.3.3 设计准则

本小节将讨论智能打字指环的设计准则，即对本技术的交互设计细节做出解释。以下讨论可能会涉及到实验二的结论，该实验收集并分析了用户的打字数据，读者可以在后面的第 4.5 节找到有关实验二的详细信息。

(1) 为什么使用长按、滑动等触摸手势选词和删除单词，而不使用按钮？：在智能打字指环的键盘布局中，26 个键已经占满了整个食指的可触及区域，因此，没有额外的空间来安放进入用于选词的数字按钮和删除单词按钮。首先，这一设计降低了用户在定位所想字母时所需要的认知负担，例如，因为在 26 键的上方没有数字按钮，所以用户在输入键盘中第一行的字母时，只需将手指尽可能伸直，而无需担心手指点中了数字按钮；其次，这一设计提高了单词联想的准确率，因为先前

工作已经证明，小屏幕文本输入中键盘布局越大，单词联系的准确率也越大^[9]。

(2) 为什么可选的候选词是五个?: 一方面，如果候选词较少，则用户想要的词很有可能未出现在候选词列表中。根据实验二的模拟实验结果，单词解码器将用户所需单词排在候选词前五名的概率为 97.5%，而排在前三名的概率仅为 93%。如果用户所需单词不在候选词列表中，用户只能重新输入，影响打字效率。另一方面，如果候选词较多，用户在选词过程中，可能会因为对手指指向控制不够准确而选错单词。综合上述原因，五个候选词是交互效率最高的选择。

(3) 为什么选词过程中，手指旋转 7.5° 对应光标移动一格候选词?: 实验二收集了用户使用智能打字指环打字时的数据。数据分析显示，被试手腕舒适的左右旋转角度在 30° 左右，因此，智能打字指环将候选词区域中的五个格子映射到-15°、-7.5°、0°、7.5° 和 15° 的手指左右选择角度上。

(4) 为什么将智能指环佩戴在食指第二指骨上? 为什么要求用户在打字的时候手腕不能平移?: 首先，这两条约束性的交互设计对于智能打字指环的单词联想来说是必须的，智能打字指环的工作原理是将惯性传感指环的方位角（偏航角和俯仰角）映射到键盘布局的 X 轴和 Y 轴上的。如果将智能指环佩戴在食指第一指骨上，用户打字过程中无论手指点击哪一行，智能打字指环的俯仰角都不会发生很大的变化，因此指环信号不能提供用户点击了哪一行字母的有效信息；如果用户在打字过程中通过手腕左右平移来点击不同列上的字母，则智能打字指环的偏航角不会发生很大变化，无法提供用户点击了哪一列字母的有效信息。其次，这两条约束性的交互设计对用户的主观体验而言尚可接受。上一章的指环佩戴位置主观喜好程度调研发现，智能指环最被用户接受的佩戴位置是手指的第一指骨，其次第二好的就是第二指骨了。手腕不允许平移的设计也是可以接受的，本章的实验二发现，手腕左右旋转在 30° 的范围内是舒适的。

4.4 指环上的触摸手势识别

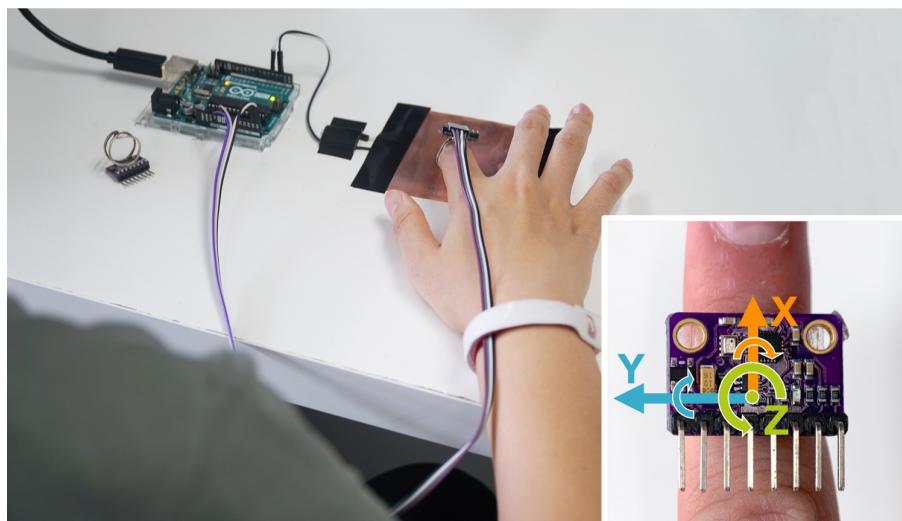
4.4.1 实验一：收集触摸手势数据

实验一收集被试佩戴惯性传感指环在桌面上触摸交互的数据，包括点击、长按、滑动等交互任务。实验的动机是为触摸手势识别技术的研发提供数据支持。在此，本节先对手指的触摸事件和抬起事件给出明确的定义：**触摸事件**指的是触摸交互中，手指接触到交互表面的一瞬间。在手指接触到交互表面以后，手指可能立即抬起，也可能在停留等待或在平面上移动过后再抬起。**抬起事件**指的是手指触摸交互表面后，离开交互表面的瞬间。本节将分三个步骤探索触摸手势的识别问题：

- (1) 对于**触摸事件**，其识别方法已经在上一章详细介绍了，本节将简要回顾。
- (2) 对于**抬起事件**，本节将结合触摸运动模型，设计抬起事件的识别算法。
- (3) 基于对触摸事件和手起事件的识别，本节采用机器学习支持了对**长按、左滑、右滑等触摸手势**的识别。

4.4.1.1 实验设计和过程

实验者从大学校园中招募了 12 名被试参与实验，被试的年龄从 19 岁到 25 岁不等，平均年龄为 21.59 岁，标准差为 2.02 岁，其中 4 名被试为女性。所有被试都是右撇子。如图4.3所示实验一的实验设置，在木质的桌子上固定着一块薄而坚硬的触摸板。被试将惯性传感指环佩戴在食指的第二指骨上，并在触摸板上执行实验要求的各种触摸手势。实验任务是实验者口头传达的，共包含四组数据集的收集：点击、长按、滑动和空中手势。其中，点击、长按和滑动用于采样各种触摸手势下的触摸事件和抬起事件的信号样本，而空中手势则作为触摸事件的负样本。



图中展示了实验一的实验设置，小图展示了惯性传感指环的佩戴位置和指环的坐标轴。

图 4.3 实验一的实验设置

(1) 点击数据集：被试在触摸板上点击 5 组 $\times 100$ 次 = 500 次。被试以自己喜欢的方式以任意的触点位置、触摸力度和角度进行点击，实验共收集了来自 12 名用户的 6000 次点击样本。

(2) 长按数据集：被试在触摸板上长按 5 组 $\times 100$ 次 = 500 次。在长按过程中，被试可以自行选择将手指保持不动，或者在交互表面上移动手指。

(3) 滑动手势数据集：被试共执行了 10 组 $\times 100$ 次 = 1000 次滑动手势，包括 5 组左滑和 5 组右滑。实验者要求被试想象自己像在智能手机上切换屏幕一样去执行滑动手势。在收集点击数据集、长按数据集和滑动手势数据集的过程中，被试需要将手

腕休息在桌子上，不能平移手腕，这一要求与智能打字指环的交互流程保持一致。

(4) 空中手势数据集：被试在一分钟的时间内，根据实验者的口头指令，在空中做出五种空中手势。手势分别是在空中保持不动、在空中左右滑动手指、画圆、画正方形和手指剧烈颤抖。这些空中手势是触摸事件的负样本，特别是手指剧烈颤抖的信号，是波形最接近触摸事件的空中手势信号，这说明本数据集是非常严格的。

在每两组实验之间，被试被要求休息一分钟，以避免疲劳。本实验的总时长大约为一小时。

4.4.1.2 实验设备

本实验的实验设备如图4.3所示。其中，惯性传感指环由 GY-91 惯性传感器和普通的装饰性指环粘合而成。为了让智能指环适合不同手指粗细的用户，实验者共制作了四枚大小不一的惯性传感指环，供用户选择。惯性传感单元通过杜邦线连接到 Arduino (Uno R3) 开发板上。实验者用魔术贴将杜邦线固定在被试的手腕上。该指环收集六轴的运动信号，包括原始加速度和角速度。其中，原始的加速度信号是线性加速度和重力混杂的结果，系统通过 Madgwick 滤波器^[82]将原始加速度分解为线性加速度和重力。实验中并未收集惯性传感指环的磁力计数据，这是因为磁力计的采样频率太低，而且容易受到周围设备的影响。图4.3中的触摸板是一块连接到 Arduino 开发板上的导电铜片。当被试的手指接触到导电铜片的时候，手指与铜片形成耦合电容，电容增大，该实验系统利用这一现象来判断手指是否接触到交互表面^[89]，作为实验的真值。以上实验设置与上一章中的数据收集实验的设置很类似，但数据收集的性能得到了很大的改进，惯性传感指环和触摸板的传感频率都达到了 1000 赫兹，触摸板的识别延迟低于 1 毫秒。这些数据采集性能的提升来自采样代码的优化、实验主机性能的提高和交互表面的小型化。最终，实验共收集了 11 个维度上的运动数据，包括时间戳、三轴线性加速度、三轴角速度、三轴重力（方向）和触摸事件的真值。

4.4.2 触摸手势识别技术

4.4.2.1 触摸检测

触摸事件指的是触摸交互中，手指接触到交互表面的一瞬间。虽然上一章已经详细介绍了基于惯性传感指环的触摸事件识别技术，但是本章仍然回顾并在新的数据集上重新评测该技术。这是因为本章所采集的数据与上一章有所不同，更具有多样性（包含各种触摸手势）、采样频率更高、无位移信号、惯性传感指环佩

戴在食指的第二指骨上，这些都是上一章没有评测过的情况。

实验者沿用第3中章低延迟触摸识别技术的代码，修改程序使之适应于 1000 赫兹的惯性传感信号，然后用实验一收集的数据对该技术进行评测。结果表明，触摸识别技术在本数据集下的召回率为 99.37% ($SD=0.94\%$)，精准率为 99.24% ($SD=1.03\%$)，平均延迟为 14.04 毫秒 ($SD=7.40$)。该结果在准确率上与上一章结果（表3.3）相比差异不大，这是因为本实验所采集的数据采样频率更高，但没有基于视觉的手指位移数据，这两点因素一正一反地影响了触摸事件识别的准确率。该结果与上一章结果相比延迟多了 5 毫秒，这一差异不是技术实现造成的，而是因为本实验的实验数据包含了左滑和右滑的数据。在这两种触摸手势下，由于用户在组织手指运动时，有部分的力量花在了手指运动的水平分量上，使得在手指运动在垂直方向上不如直接的点击运动快速，根据触摸运动模型对触摸识别技术的指导意义（第 3.8.1 小节），这种情况下的识别延迟的确会增大。在误触率方面，该技术在 12 名被试共 60 分钟的空中手势数据集中仅触发了 14 次误触事件，这说明，即使用户在空中做出诸如剧烈抖动手指等接近触摸碰撞的空中手势，该触摸识别技术依然不会轻易引发误触。

4.4.2.2 触摸与抬起的相关性

(1) 大多数情况下，触摸发生在触摸运动中点 $\frac{t_1}{2}$ 之前，抬起发生在中点之后。

如图2.10所示，触摸运动模型蕴含了触摸时间 t_c 、触摸运动中点 $\frac{t_1}{2}$ 和抬起时间 t_{up} 的关系：

$$t_c < \frac{t_1}{2} < t_{up} \quad (4.1)$$

其中，触摸运动中点 $\frac{t_1}{2}$ 是触摸运动方程总时长的一半，也是人对手指的控制从向下加速到向上减速变化的瞬间：

$$\begin{cases} a(t) < 0 & t < \frac{t_1}{2} \\ a(t) \geq 0 & t \geq \frac{t_1}{2} \end{cases} \quad (4.2)$$

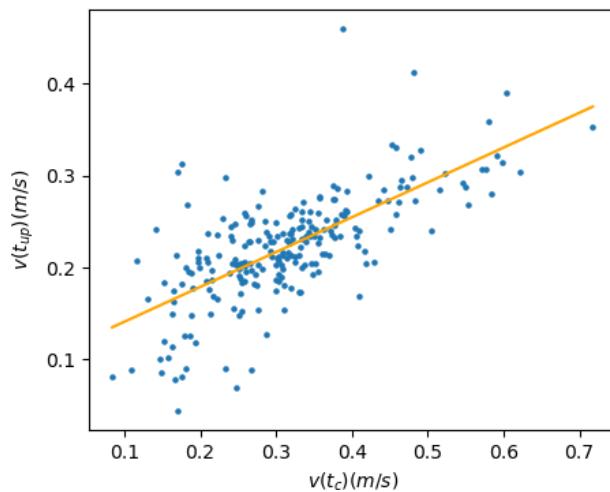
根据第2章的实验数据，在点击和滑动等短时触摸交互中， $t_c < \frac{t_1}{2}$ 成立的比例为 92.86%。也就是说，大多数触摸运动都是一个手指向下加速过程，在手指还未开始减速时，手指已经因为接触交互表面而瞬停。

根据第2章的实验数据，在点击和滑动等短时触摸交互中， $\frac{t_1}{2} < t_{up}$ 成立的比例为 96.45%。这一现象可由触摸运动模型解释：当 $t < \frac{t_1}{2}$ 时，人仍然对手指施加着

一个向下的力，此时即使手指已经因为交互表面的阻挡而急剧减速，但尚不会反向运动离开交互表面；而当 $t > \frac{t_1}{2}$ 时，人对手指施加的力转向向上，随后手指的运动才可能转向向上。这一规律对触摸抬起事件的识别有很大帮助，本章中后续介绍的实验证明，该判据有效降低了抬起事件识别的误报率。

(2) 触摸速度与 $v(t_c)$ 与抬起速度 $v(t_{up})$ 之间存在正相关性。

4.4.2.3 触摸与抬起的运动速度



触摸与抬起的运动速度之间存在着正相关关系，因此触摸速度可以作为抬起事件识别的辅助判据。

图 4.4 触摸与抬起速度之间的关系

对于点击和滑动等短时触摸交互而言，手指向下点击的速度与抬起速度之间存在正相关的关系。如图4.4所示，根据第2章收集的实验数据，触摸瞬间手指的速度 $v(t_c)$ 与抬起瞬间手指的速度 $v(t_{up})$ 之间存在关系：

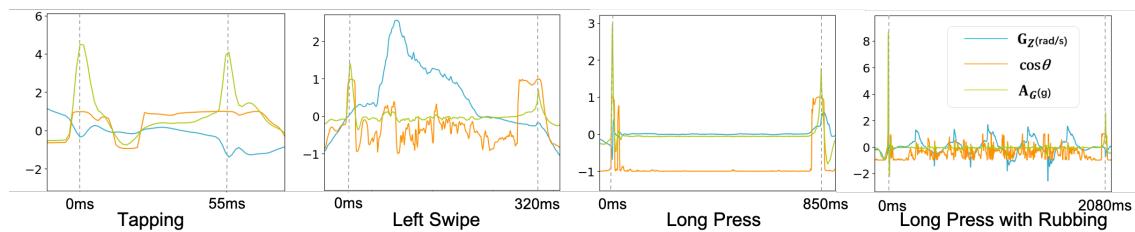
$$v(t_{up}) = 0.38 \times v(t_c) + 0.10 \quad (4.3)$$

其中，线性拟合的决定系数 (R2) 为 0.48，说明 $v(t_c)$ 与 $v(t_{up})$ 之间存在弱相关性。因此，触摸速度这一信息可以作为抬起事件识别的一个辅助判据。

为了验证本文模型对触摸手势识别的优化是否实际有效，本章将基于上述优化思路，实现一种指环上的高效打字技术，该技术涉及到点击、长按、左滑和右滑等多种触摸手势。文本输入作为最快速、最复杂的触摸交互任务之一，可以严格地验证本章触摸识别技术的可用性。

4.4.2.4 抬起事件识别

抬起事件指的是手指触摸交互表面之后，离开交互表面的瞬间。抬起事件识别技术不仅要准确的、高响应地识别手指离开交互表面的瞬间，还要防止在手指接触着交互表面时，将手指在表面上的移动误识别成抬起事件。虽然触摸识别模型不包含对手指抬起过程的描述，但是该模型仍然对抬起事件的识别具有指导意义。



图中展示了四种触摸手势下，从手指开始下落到最终离开交互表面的时间内，惯性传感器的相关信号。四种触摸手势分别是点击、左滑、长按和长按并晃动。

图 4.5 触摸交互信号图示

如图4.5所示是一次完整的触摸交互的相关信号，从左到右四幅子图分别是点击、左滑、长按和长按并晃动手指这四种情况下的惯性传感器相关信号，每幅子图中有两条竖虚线，表示触摸事件和抬起事件的发生时间。图中的三条曲线并不是指环上惯性传感器的原始信号波形，而是经过了提炼的，能够较好地检测抬起事件的特征值的曲线。绿线 A_G 是线性加速度在重力方向上的投影，相当于触摸运动模型中，手指与交互表面距离 x 对时间的二阶导。 A_G 在触摸事件和抬起事件发生的两个时间点上的振幅很大，波峰也十分尖锐。橙线 $\cos\theta$ 是加速度方向和重力方向的夹角的余弦值，蓝线 G_z 是手指方位角上的角速度。从图中的波形可以看出，上述三个值都在手机抬起时具有明显的特征，可用于识别抬起事件。

然而，仅根据惯性传感器信号和上述指标识别抬起事件，准确率仍然较低。为了提高抬起事件的识别准确率，此处引入本章引言中提到的基于触摸运动模型的两条优化思路：

(1) 大多数情况下，抬起事件出现在触摸运动中点之后，即 $t_{up} > \frac{t_1}{2}$ 。这一规律可以有效处理指环上的抬起事件识别所面临的一个现实工程问题，即惯性传感器在遭到强烈震动（如手指接触交互表面）后，其信号会有 10 到 30 毫秒的时间处于紊乱状态。实验者推测，该紊乱状态是由惯性传感器自身原理和 Madgwick 滤波器共同造成的，难以避免。在这短暂的紊乱时间内，其信号很容易被误识别为抬起事件。根据公式 $t_{up} > \frac{t_1}{2}$ ，实验者将当前时间 t 和触摸运动中点 $\frac{t_1}{2}$ 这两个值也加入到识别抬起事件的特征组之中，如此一来，只要触摸运动的进度未过半，惯性传感器的紊乱就不太可能导致抬起事件的误报。

(2) 触摸速度 $v(t_c)$ 与抬起速度 $v(t_{up})$ 之间存在正相关的关系。根据这一规律，实验者将手指触摸瞬间的运动速度 $v(t_c)$ 也加入到识别抬起事件的特征组之中，如此一来，重触带来的大幅度信号紊乱更不可能引发抬起事件的误报，而轻触之后的抬起事件识别也能更为敏锐。

综上所述，本章采用触摸运动模型与机器学习相结合的方式来识别抬起事件：对 50 毫秒内的信号时间序列，提取 21 维特征向量，包含 A_G 、 $\cos\theta$ 、 G_z 的最小值、最大值、均值、偏度和峰度，外加 t 、 $\frac{t_1}{2}$ 和 $v(t_c)$ 这三个值。其中， t 、 $\frac{t_1}{2}$ 和 $v(t_c)$ 可以通过第二章所述的触摸运动模型的计算方法获取。实验者从点击、长按和滑动数据集中采样训练数据，正样本来自触摸手指抬起时刻周围的时间窗口 [-25ms, 25ms]，共 24000 份正样本。由于抬起事件识别的误触发指的是错把手指接触着交互表面时的动作误识别为抬起事件，因此负样本从手指接触着交互表面时的数据中随机截取。实验者从每一次的触摸数据中随机切取三段 50 毫秒的信号作为负样本，总共收集到 54000 个负样本。最后，基于上述正负样本数据，实验者采用简单的机器学习方法支持向量机（SVM）训练了一个二分类模型，模型的输入时任意 50 毫秒的 21 维特征向量，输出是这 50 毫秒的结束时刻是否发生了抬起事件。

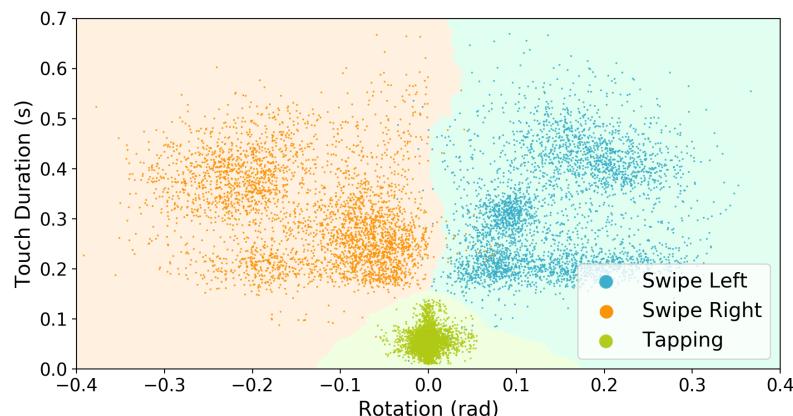
留一（被试）交叉验证显示，模型在点击、长按和滑动数据集上预测抬起事件，召回率分别为 98.61% ($SD=1.87\%$)、99.68% ($SD=1.02\%$) 和 99.53% ($SD=1.39\%$)，精准率分别为 99.82% ($SD=0.24\%$)、99.78% ($SD=0.31\%$) 和 99.65% ($SD=0.42\%$)，识别延迟分别为 8.55 毫秒 ($SD=4.97$)、7.83 毫秒 ($SD=3.47$) 和 7.97 毫秒 ($SD=3.77$)。结果表明，上述结合了机器学习和触摸运动模型的抬起事件识别模型，在预测准确率和延迟上都达到了较高的水平。

4.4.2.5 长按

基于触摸事件和抬起事件的识别，实验者采用阈值方法识别长按手势：当用户的手指接触交互表面并停留超过 200 毫秒时，系统报告长按事件。其中，停留值得是没有发生手指抬起事件，且手指上惯性传感指环的角速度始终小于 0.1 rad/s。角速度的阈值是一个经验值，它确保大多数滑动手势不会被误识别成长按。为了评测上述长按识别方法的性能，实验者从长按数据集中采样正样本，从点击和滑动数据集中采样负样本。最终，留一（被试）验证法显示，该识别方法的召回率为 98.8%，精准率为 99.2%。后续的文本输入实验也进一步说明，本长按识别方法在实用中不存在任何问题。

4.4.2.6 左滑和右滑

由于长按事件已经被准确识别，所以接下来只需要区分点击、左滑和右滑三种触摸手势即可。实验者设计了两个特征来对上述三种触摸手势进行分类：触摸持续时长和旋转角度。图4.6对比了点击、左滑和右滑三种触摸手势下，触摸持续时长和旋转角度的差别。从图中可以看出，三种触摸手势在上述两个特征的笛卡尔积平面上的点云是相互远离的，因此，一个简单的KNN（ $K=10$ ）模型就可以准确区分这三种触摸手势。留一（被试）交叉验证法显示，该模型的三分类准确率为99.38%（ $SD=0.43\%$ ）。在实际使用中，系统将在检测到抬起事件时调用此三分类模型，判断用户是否执行了左滑或右滑触摸手势。



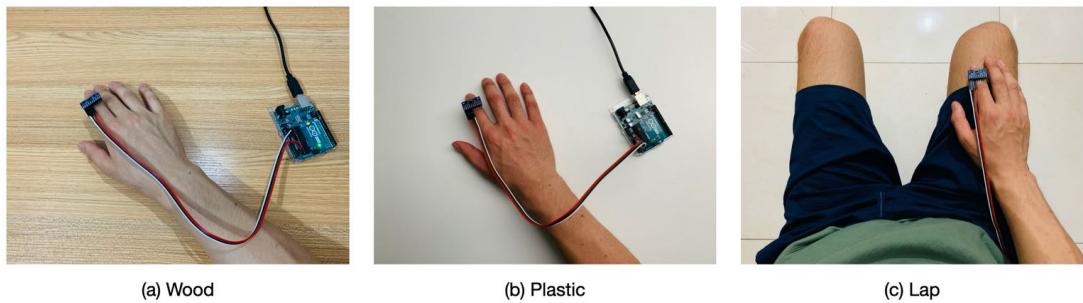
如图所示是点击、左滑、右滑三种触摸手势在触摸持续时长和旋转角度这两个特征的笛卡尔积表面上的点云分布图，从图中不难发现，基于触摸持续时长和旋转角度这两个特征的KNN模型将能准确分类点击、左滑和右滑。

图 4.6 点击、左滑、右滑三分类模型

4.4.2.7 不同表面的触摸手势识别

上述研究评测了触摸手势识别技术在木质桌面上的表现。在之前的数据采集实验中，实验者收集了被试佩戴运动传感指环在木质桌子上的触摸交互数据，桌子上还附有一块薄铜片，作为收集触摸与否的真值的触摸板。然而，之前的实验方法无法完美还原用户在无源表面上触摸交互的情形，因此，实验者进行了一项额外的实验，以在不同的无源表面上评测触摸手势识别算法，无源表面包括木质桌面（无薄铜片）、塑料桌面和人的大腿上侧（如图4.7所示）。

共有12名被试参与了实验，被试的年龄从20岁到26岁不等，平均年龄为22.03岁，标准差3.37，其中3名被试为女性。他们没有参加过之前的用户实验。本实验的实验设置与上一个实验基本相同，只有以下不同点：（1）在三种类型的无源表面开展实验，手机点击、长按和滑动数据集；（2）交互表面无薄铜片（触摸



本实验评测了触摸手势识别技术在不同无源表面上的表现，无源表面包括：(a) 木质桌面，(b) 塑料桌面，和(c) 人的大腿上侧。

图 4.7 实验设置

板)附着；(3)实验持续了三天，每天一个小时，实验分别在三个不同的无源表面上进行；(4)被试在每触摸 5 次之后，需要按回车键以继续实验。

本实验的挑战在于我们无法通过触摸板来获得触摸与否的真值。为了克服这一问题，实验者根据波形和实验录像人工标注了触摸事件和抬起事件。由于被试每五次触摸就会按回车键确认一次，因此可以通过触摸的数量是否对得上号来保证人工标注的正确性。实验者使用三个表面的所有数据来训练通用化模型，使这一个模型能适用于不同的无源表面。最后，实验者采用留一(被试)交叉验证来评测模型的性能。

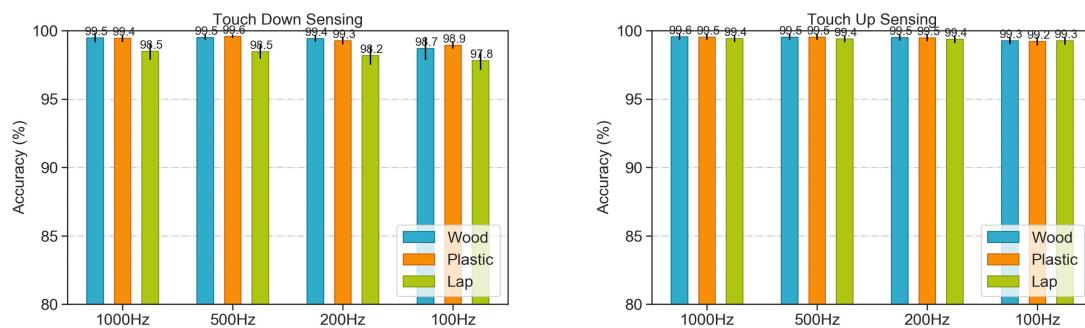
	触摸事件		抬起事件	
	召回率	精准率	召回率	精准率
木板	99.54% (0.82%)	99.45% (0.46%)	99.45% (0.58%)	99.65% (0.23%)
塑料	99.67% (0.82%)	99.22% (0.46%)	99.38% (0.68%)	99.71% (0.23%)
大腿	97.69% (1.42%)	99.32% (0.76%)	99.18% (0.68%)	99.66% (0.15%)

表 4.1 触摸手势识别技术在不同无源表面上的表现

表4.1展示了评测结果。对于触摸事件识别，重复测量方差分析显示，交互表面对召回率有显著影响 ($F_{2,22} = 25.14, p < .001$)。Bonferroni 校正后的后验测试显示以下交互表面之间存在显著差异：木板-大腿 ($p < .005$) 和塑料-大腿 ($p < .005$)。也就是说，触摸事件识别方法在刚体表面（如木质桌面和塑料桌面）上表现良好，但当用户在大腿上进行触摸交互时，该技术存在一些漏识别的情况。对于抬起事件识别，交互表面对召回率和精准率都没有显著性影响，这说明，抬起事件识别方法在所有的测试表面上都表现良好。在实验后，实验者让被试简单试用该触摸事件识别技术，被试们发现，即使他们故意很轻地去触摸，该技术还是能准确地汇报触摸事件。

4.4.2.8 降采样

以上实验评测了 1000 赫兹惯性传感信号下触摸手势识别的性能。然而，实际的交互系统中可能通过频率较低的惯性传感器来节约设备成本和电能。由于惯性传感器的频率降低会导致识别准确率也降低，实验者通过降采样的模拟实验来评估识别方法在 500 赫兹、200 赫兹和 100 赫兹的传感频率下的准确率，结果如图4.8所示，其中，左图是触摸事件识别的准确率，右图是抬起事件识别的准确率，准确率指的是 F1 综合评价指标。



如图所示，当采样频率降低时，触摸事件和抬起事件的识别准确率都会受到影响，但总体而言，200 赫兹下的识别准确率已经很高。

图 4.8 降采样情况下触摸手势识别的准确率

对于触摸事件识别，双因素重复测量方差分析显示，传感器频率对识别准确率有显著影响 ($F_{3,33} = 16.44, p < .001$)。Bonferroni 校正的后验测试显示以下频率之间存在显著差异：1000Hz-100Hz ($p < .001$)、500Hz-100Hz ($p < .001$)、200Hz-100Hz ($p < .05$)。对于抬起事件识别，频率对性能也有显著影响 ($F_{3,33} = 39.94, p < .001$)，Bonferroni 校正的后验测试显示以下频率之间存在显著差异：1000Hz-100Hz ($p < .001$)、500Hz-100Hz ($p < .001$)、200Hz-100Hz ($p < .001$)。结果表明，虽然降低频率的确会影响触摸手势识别的准确率，但是 200 赫兹以上的传感信号已经能提供较高的识别准确率了。

4.5 智能打字指环的解码器设计

4.5.1 实验二：收集打字数据

本实验的目的是收集被试的打字行为数据，用以设计文本输入单词解码器。在使用智能打字指环输入文本时，用户不能在无源表面上看到键盘布局，相反，用户需要回忆平时打字时每个字母的键位，在桌面上想象一个键盘来打字。本实验通过收集用户打字的数据来确定想象键盘的空间位置参数，实验者也通过这个实验

来观察被试的打字行为。

4.5.1.1 实验设计和过程

实验者从校园中招募了 12 名被试，被试的年龄从 19 岁到 29 岁不等，平均年龄为 23.67 岁，标准差为 3.14，其中 4 名被试为女性。以上被试均为参与过之前的用户实验。在正式实验开始前，有一个简短的熟悉阶段。实验者向被试介绍想象键盘的概念，特别是向被试介绍清楚想象键盘的边界（如图4.1B-E 所示）。然后，被试按顺序键入从 A 到 Z 的二十六个英文字母，以熟悉自己心中想象键盘中每个按键的位置。通过平移手腕而不是旋转手腕来打字是智能打字指环的用法中禁止的，但时常被被试忽略，因此，实验者在熟悉阶段中会纠正被试平移手腕的问题。熟悉阶段的时长大约为五分钟，实验由五段重复的实验组成。在每段实验中，被试需要眷写十个短句，这些短句是从 MacKenzie 短句集^[151]中随机抽取的。



图中展示了实验二的实验设置，被试佩戴惯性传感指环在一张普通桌子上打字，通过一个显示器接收视觉反馈。

图 4.9 实验二的实验设置

如图4.9所示，被试坐在可调节的座椅上进行实验。被试可以将椅子调整到最舒适的高度和角度。一台显示器中展示了用户界面，包括二十六键键盘的布局、任务中要求眷写的句子，和已经输入的单词。被试佩戴智能打字指环，在一张普通的桌子上打字，实验者要求被试尽可能又快又准地眷写短句。在实验二中，系统还没有单词解码器，用户不能真的按自己的意愿输入单词：无论被试如何打字，系统都将始终显示正确的字符。但是，如果被试自己主观上发现自己打错了，则应

该重做正在输入的短句。之前也有工作^[98,150,152]采用相同的方法来收集理想的打字行为数据。实验中，被试被要求将视觉注意力集中在显示器上，而不是看着自己的手部。整个实验的时长约为一个小时。

4.5.1.2 数据处理

智能打字指环的基本思想是通过惯性传感指环的方位角来预测用户想要输入的单词。惯性传感器的俯仰角和偏航角对单词解码器来说有用：(1) **俯仰角 (Pitch)** 指的是手指指向与水平面的夹角，当用户在想象键盘上点击位于不同行的字母时，惯性传感指环的俯仰角也是不同的；(2) **偏航角 (Yaw)** 指的是手指绕竖直轴的旋转角度，当用户在想象键盘上点击位于不同列的字母时，偏航角是不同的。实验者使用 Madgwick 滤波器^[82]来获取惯性传感器的俯仰角。然而，惯性传感器所提供的信息无法获取绝对的偏航角，实验者通过偏航角上角速度的积分来估计两次点击之间的偏航角增量 (ΔYaw)：

$$\Delta Yaw = \sum_{i=1}^T G_{zi} \Delta t_i \quad (4.4)$$

其中， T 是连续两次点击之间的信号帧数， G_{zi} 是第 i 帧中惯性传感器角速度的 z 轴分量， Δt_i 是第 i 帧的持续时长。

接下来，首先形式化本节可能涉及的量。假设实验中，被试们点击字母 u 共计 N_u 次，本文将所有敲击字母 u 时惯性传感指环的俯仰角的集合表示为 $P_u = \{Pitch_{ui}\}_{i=1}^{N_u}$ ，将偏航角的集合表示为 $Y_u = \{Yaw_{ui}\}_{i=1}^{N_u}$ 。假设实验中，被试们共有 $N_{u,v}$ 连续点击了 u 和 v 这两个字母，本文使用 $\Delta Y_{u,v} = \{\Delta Yaw_{u,vi}\}_{i=1}^{N_{u,v}}$ 来表示它们的偏航角增量的集合。受限于惯性传感器的能力，本实验所收集的数据包括绝对俯仰角 P_u 和相对偏航角 $\Delta Y_{u,v}$ ，但不包括绝对偏航角 Y_u ，这是因为，六轴的惯性传感数据不可能还原出绝对偏航角。对于每个被试的 P_u 和 $\Delta Y_{u,v}$ ，实验者都剔除了三个标准差以外的极端数据。

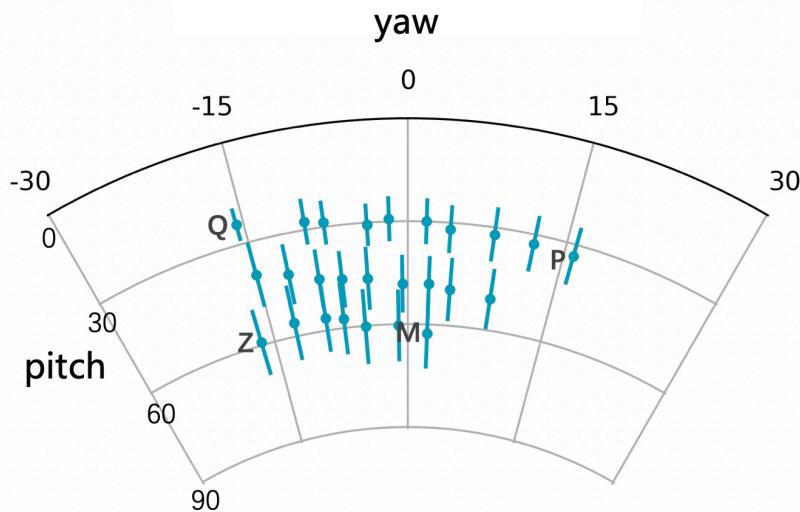
4.5.1.3 实验结果

同一个用户想要键入同一个字母 u 时，其手指的方位角是类似的。因此， P_u 和 Y_u 应服从某种概率分布。根据论文 BlindType^[98] 中所介绍的相对触点模型， $\Delta Y_{u,v}$ 也服从某种概率分布。为了简化计算，实验者假设 P_u 和 $\Delta Y_{u,v}$ 都符合二维正态分布，这次文本输入相关工作中常用的假设^[94,98]。由于 P_u 和 $\Delta Y_{u,v}$ 是由实验直接测得的，实验者拟合了它们的二维正态分布 $(\bar{P}_u, \sigma P_u)$ 和 $(\bar{Y}_{u,v}, \sigma Y_{u,v})$ ，单词解码器会用到以上两个分布，这将在后续的一节中介绍。然而， Y_u 不可由实验设备直接测

得，实验者间接地通过求解以下最优化问题来估计它的取值，使之最符合实验的观测结果：

$$\begin{cases} \bar{Y}_u \\ \min \sum_u \sum_v N_{u,v} (\bar{Y}_v - \bar{Y}_u) - \Delta Y_{u,v}^2 \end{cases} \quad (4.5)$$

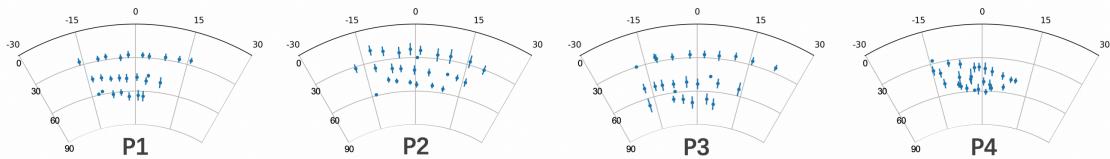
其中， \bar{Y}_u 是点云 Y_u 的均值，实验者通过最速下降算法求解上述最优化问题。



如图所示是被试们使用智能打字指环打字时，二十六个键位对应的俯仰角和偏航角点云。其中，误差条表示俯仰角的标准差。

图 4.10 平均意义下的触点点云

如图4.10所示是所有被试平均情况下的二十六键触点点云 (\bar{Y}_u, \bar{P}_u) ，它反映了平均意义下被试们心中想象键盘的布局。如实验者所期待，惯性传感器的俯仰角和偏航角很好地表征了用户所想输入按键的行号和列号。想象键盘的左边边界在字母 Z 上，右边边界在字母 P 上，其偏航角跨度为 32.4° 。想象键盘的上下边界在字母 Q 和 M 上，分别对应 24.1° 和 62.6° 俯仰角，其跨度为 38.4° 。



如图所示是四名不同被试的触点点云，误差条表示俯仰角的标准差。

图 4.11 不同被试的触点点云

图4.11展示了四名典型被试的二十六键触点点云，从图中可以看出，被试们心中的想象键盘在大小和形状上都存在较大差异。最大的想象键盘（被试 P3）的偏航

角跨度和俯仰角跨度为 $40.1^\circ \times 45.7^\circ$, 而最小的想象键盘(被试 P4)为 $25.3^\circ \times 33.4^\circ$ 。被试 P1 通常以差不多的俯仰角点击位于键盘同一行的按键, 也就是说, 他的想象键盘是扇形的。被试 P2 在点击左上角的按键时手指伸得更远, 实验者推测, 这位被试的想象键盘是一个向左偏的矩形, 因此他认为左上角的键是最远的。

4.5.2 智能打字指环单词解码器

智能打字指环单词解码器的工作原理是贝叶斯方法, 该方法通过以下公式评估单词 W 作为用户所想单词的概率, 然后选出概率最大的五个候选词供用户选择:

$$P(W|I) \propto P(I|W) \times P(W) \quad (4.6)$$

其中, I 是用户输入此单词时的触点序列, 每个触点用手指接触交互表面时惯性传感器的偏航角和俯仰角表示。本节将从触点模型 $P(I|W)$ 和语言模型 $P(W)$ 两个方面介绍单词解码器。本节中涉及到的模拟实验评测都基于实验二中收集的数据。

4.5.2.1 触点模型

$P(I|W)$ 是贝叶斯单词解码器的触点模型部分:

$$P(I|W) = \prod_{i=1}^n P(I_i|W_i) \quad (4.7)$$

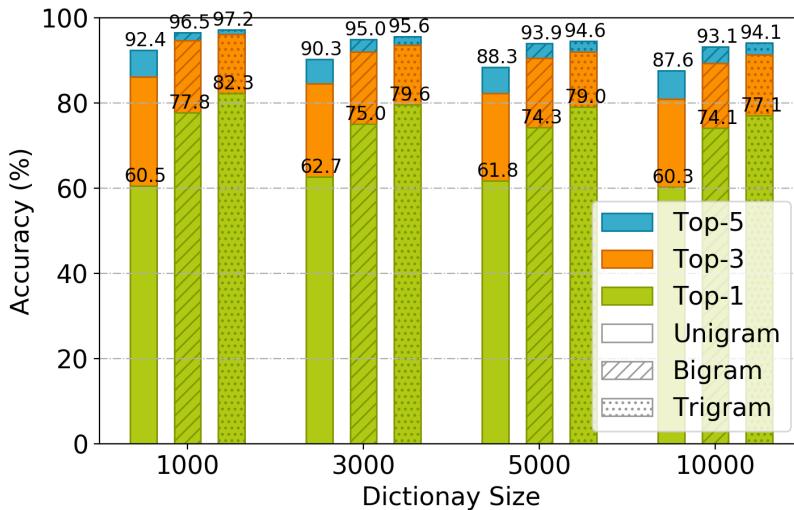
其中, n 是单词 W 的长度, W_i 表示单词的第 i 个字母, I_i 表示用户输入该单词时的第 i 次触点。在类似的工作中, 研究者一般认为触点服从二维高斯分布^[148,150], 这是绝对触点模型; 同时, 也有研究者指出, 连续两次触点之间的向量也服从二维高斯分布^[98], 这是相对触点模型。在本节所介绍的单词解码器中, 受限于惯性传感指环的识别能力, 实验者假设绝对触点模型适用于触点的俯仰角, 而相对触点模型适用于偏航角。也就是说, 触点俯仰角集合 P_u 和相邻两次触点之间的偏航角增量集合 $\Delta Y_{u,v}$ 都服从二维正态分布, 因此, $P(I|W)$ 可以表示为:

$$P(I|W) = \prod_{i=1}^n P(Pitch_i|W_i) \prod_{i=1}^{n-1} P(\Delta Yaw_{i,i+1}|W_{i,i+1}) \quad (4.8)$$

其中, $Pitch_i$ 是用户输入字母 W_i 时惯性传感指环的俯仰角, $\Delta Yaw_{i,i+1}$ 是用户在输入字母 W_i 和 W_{i+1} 的时间段内手指的偏航角增量。单词解码器可以利用

实验二中整理出来的分布 P_{W_i} 和 $Y_{W_i, W_{i+1}}$ 来计算上述公式中的值 $P(Pitch_i | W_i)$ 和 $P(\Delta Yaw_{i,i+1} | W_{i,i+1})$ 。

4.5.2.2 语言模型



如图所示是不同语言模型下，解码器在不同词库大小下的单词预测准确率。

图 4.12 不同语言模型下解码器的预测准确率

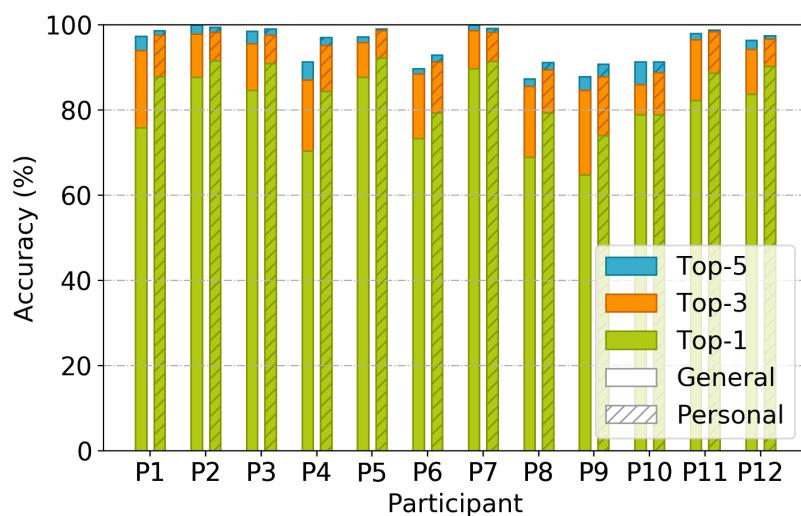
$P(W)$ 是单词解码器中的语言模型部分。实验者通过模拟实验测试了单元、双元、三元语言模型的性能^[149]。单元语言模型通过大型预料库中每个单词的出现频次来估计每个单词出现的概率 $P(W)$ ，双元语言模型中的 $P(W)$ 是将前一个已输入单词作为前提条件时下个单词为 W 的概率，三元语言模型则是考虑前两个已输入的单词。如图4.12所示是应用不同语言模型时，解码器在不同词库大小下的单词预测准确率。其中，Top-1、Top-3、Top-5 准确率指的是用户所想单词出现在候选词列表第一位、前三位和前五位的概率。

二元重复测量方差分析显示，语言模型对解码器的 Top-1 准确率有显著性影响 ($F_{2,22} = 86.53, p < .001$)。三元语言模型显著优于双元语言模型 ($p < .001$) 和单元语言模型 ($p < .001$)，在词库大小为 5000 的情况下它的 Top-1 准确率达到 79.0%，Top-5 准确率达到 94.6%。也就是说，当用户使用智能打字指环输入一篇词汇量在 5000 个单词以内的文章时，单词解码器将用户手指敲击桌面的信号直接转化为他想要的单词的概率是 79.0%，用户能通过选词功能选中所需单词的概率是 94.6%。由于三元语言模型的表现是最好的，在本章的剩余内容中都默认使用三元语言模型。方差分析还显示，词库大小对解码器的 Top-1 准确率的影响只是一个趋势 ($F_{3,33} = 2.41, p = .085$)，还不是显著的。这是因为在本实验的测试短句中，以

及在日常打字过程中，出现低频词汇的概率是很低的。

4.5.2.3 用户个性化

实验二曾发现被试的打字行为差异很大，这启发实验者研究解码器是否应该针对用户个性化做出优化。本章将**通用化模型**定义为通过实验二所有被试的数据训练得到的单词解码器；而**个性化模型**是在通用化模型的基础上，不断通过特定用户输入数据更新自身参数的单词解码器。在实际应用场景中，所有用户刚开始只能使用通用化模型作为智能打字指环的单词解码器，个性化模型不断收集用户的打字数据，并试图在数据量足够大的时候更新某些按键的触点分布 P_u 或 $\Delta Y_{u,v}$ 。更具体地，当特定用户个人的触点集合 P_u 或 $\Delta Y_{u,v}$ 的大小大于等于 8 时，个性化的单词解码器就会根据触点集合重新计算其分布，其中，阈值 8 是模拟实验得到的最佳阈值。



图中展示了通用化模型和个性化模型在不同用户的数据上的预测准确率。

图 4.13 通用化模型和个性化模型的预测准确率

图4.13展示了通用化模型和个性化模型的准确率对比。实验者通过五折交叉验证来评测个性化模型的性能，即将实验二中每个人的四段实验数据（40句话）作为训练集，将一段实验数据（10句话）作为测试集。从图中可以看出，个性化的单词解码器的准确率更高。下一个实验将通过正式的用户实验来进一步评测智能打字指环的性能，同时比较通用化模型和个性化模型。

4.6 智能打字指环的评测

4.6.1 实验三：智能打字指环的评测实验

实验三是用于评测的用户实验，有两个目的：一是评测基于惯性传感指环的无源表面文本输入技术（智能打字指环）的打字速度和学习曲线；二是对比上一节所述的通用化模型和个性化模型。

4.6.1.1 实验设计和过程

本实验为期五天，采用组间实验设计来对比通用化模型和个性化模型。实验者从校园中招募了 16 名被试参与实验，被试的年龄从 20 岁到 27 岁不等，平均年龄为 22.86 岁，方差为 2.09，其中有 6 名女性被试。被试们没有参与过之前的任何用户实验。在每一天的实验中，被试都会誊写 30 句话，这 30 句话是从 MacKenzie 短句库^[151]中随机抽取的。

本实验的设置和实验二相似（如图4.9）。被试坐在可调节的桌椅上，将惯性传感指环佩戴在食指第二骨节上，在一张普通的桌子上输入文本。被试们通过一台显示器接收视觉反馈。实验者要求被试尽可能又快又准地完成文本输入任务，在实验中，被试被要求将视线集中在屏幕上，而不要过多地关注自己的手部。与实验二不同的是，本实验已经应用了智能打字指环文本输入技术，被试需要真正地去打字，单词打错了要删除后重打，而不会像在实验二中一样无论怎样输入都显示正确的字母。

在第一天的实验当中，所有用户都使用基于通用化模型的单词解码器进行实验。通用化模型是由实验二中所有用户的实验数据拟合而成的。从第二天开始，我们将被试分为两组，每组都有八名被试。两组被试的平均年龄分别是 22.50 岁 ($SD=1.77$) 和 23.25 岁 ($SD = 2.55$)，他们两组人在第一天实验中的平均打字速度是尽可能接近的。具体而言，实验者写了一个分组程序，随机将 16 名被试分组 1000 次，选取两组人平均打字速度最接近的一种分组方法来实际执行。第一组被试在后续四天的实验中继续使用通用化模型；第二组被试则在后续四天实验中使用个性化模型。对于使用个性化模型进行实验的每一名被试，实验者都会在每天的实验开始前，用该被试已有数据重新拟合个性化的单词解码器。

在第一天的正式实验之前，有一个简短的热身阶段。研究者向被试介绍想象键盘的概念。在实验者的指导下，被试从 A 到 Z 点击字母各两次以熟悉想象键盘。然后，被试会尝试输入五句话。整个热身阶段的时长大约为十分钟。在每天的正式实验中，被试分三段实验来誊写 30 句话，在每两段实验之间休息五分钟的时间。每天实验的时长为一个半小时，五天实验的总时长为七个半小时。

4.6.2 智能打字指环的评测结果

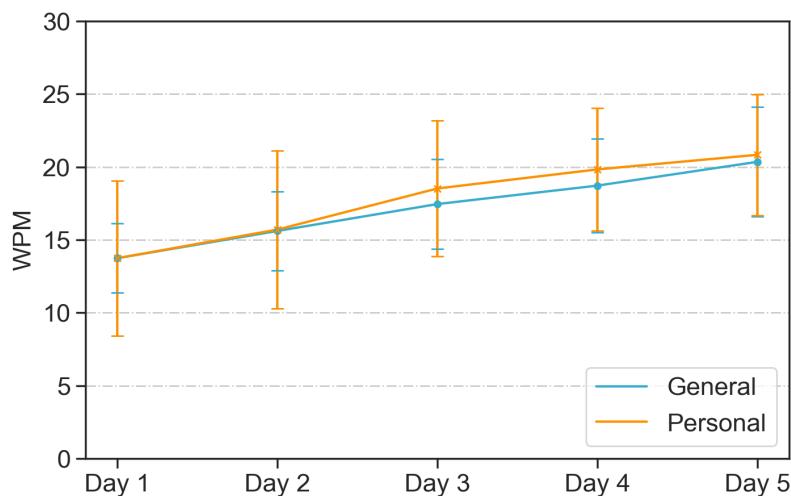
实验者使用混合方差分析来评估组间因素（解码器模型）和组内因素（实验天数）对打字速度、未纠正错误率（UER）和已纠正错误率（CER）的显著性影响。由于 UER 和 CER 不服从正态分布，实验者在方差分析之前使用对齐秩变换算法^[153]校正数据。如果有任何独立变量或组合对实验结果有显著性影响 ($p < .05$)，实验者采用 Bonferroni 校正后的后验测试来做成对比较。

4.6.2.1 打字速度

本实验对打字速度的测量单位是每分钟输入的英文单词数（WPM）^[154]，计算公式如下：

$$WPM = \frac{|S| - 1}{T} \times 60 \times \frac{1}{5} \quad (4.9)$$

其中， $|S|$ 是所眷写句子的字符长度（包括空格）， T 用户完成眷写的时长。



图中展示了被试采用通用化模型或个性化模型时的打字速度随实验天数增加而增长的曲线。

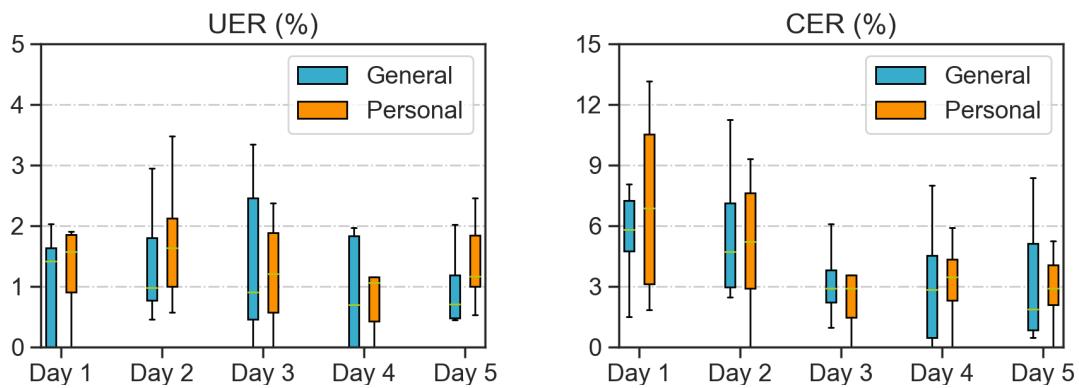
图 4.14 打字速度

如图4.14所示是五天实验中被试打字速度的变化。通用化模型下被试第一天的打字速度为 13.75 WPM ($SD=2.65$)，第五天的打字速度为 20.83 WPM ($SD=4.20$)。个性化模型下被试第一天的打字速度为 13.74 WPM ($SD=5.33$),; 第五天的打字速度为 20.83 WPM ($SD=4.14$)。模型对打字速度没有显著性影响 ($F_{1,14} = 0.09, p = 0.77$)。实验天数对通用化模型下的打字速度 ($F_{4,28} = 27.00, p < .001$) 和个性化模型下的打字速度 ($F_{4,28} = 41.17, p < .001$) 都有显著性影响。对于通用化模型，后验测试

显示以下实验天数之间存在显著差异：1-3 ($p<.05$)、1-4 ($p<.005$)、1-5 ($p<.005$)、2-4 ($p<.05$)、2-5 ($p<.05$)、3-5 ($p<.05$) 和 4-5 ($p<.05$)。对于个性化模型，以下实验天数之间存在显著性差异：1-3 ($p<.01$)、1-4 ($p<.005$)、1-5 ($p<.001$)、2-3 ($p<.005$)、2-4 ($p<.005$)、2-5 ($p<.005$)、3-5 ($p<.05$)。模型和实验天数之间不存在显著的交互作用 ($F_{4,56} = 0.60, p = 0.66$)。图中的学习曲线在最后一天中并未收敛，因此本实验不反映该技术的专家级打字速度。

4.6.2.2 打字错误率

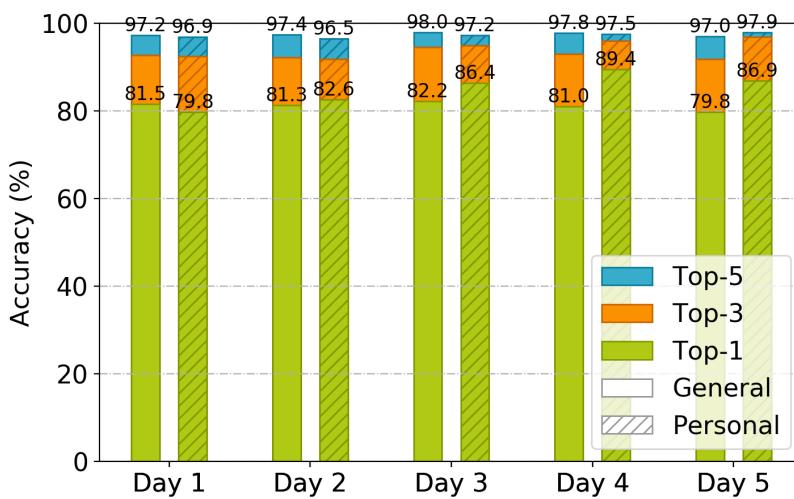
两种指标被用于评测本文本输入法的错误率：(1) 未纠正错误率 (UER) —— 遗留在所誊写文本中的错误，UER 等于未经纠正的单词数量除以所誊写句子的单词数量；(2) 已纠正错误率 (CER) —— 那些在打字过程中被纠正（例如通过删除）的错误，CER 等于被纠正的单词数量除以所誊写句子的单词数量。



图中展示了被试采用通用化模型或个性化模型时的打字错误率随实验天数而变化的箱型图。

图 4.15 错误率

如图4.15所示是五天实验中 UER 和 CER 的变化。模型和实验天数都对 UER 没有显著性影响。通用化模型下平均 UER 为 1.17% ($SD = 1.02\%$)，个性化模型下平均 UER 为 1.50% (1.40%)。模型对 CER 没有显著性影响，但实验天数对 CER 存在显著性影响 ($F_{4,56} = 6.84, p < .005$)。后验测试表明以下实验天数之间 CER 存在显著性差异：1-3 ($p<.005$)、1-4 ($p<.05$)、1-5 ($p<.005$)、2-3 ($p<.05$) 和 2-5 ($p<.05$)。在第五天的实验中，通用化模型下平均 CER 为 3.22% ($SD=2.92\%$)，个性化模型下平均 CER 为 2.92% ($SD=1.65\%$)，也就是说，被试每输入 30 个英文单词才需要纠错一次。



图中展示了两种模型下单词解码器的准确率随实验天数而变化的柱状图，准确率指标包括 Top-1、Top-3 和 Top-5 准确率。

图 4.16 单词解码器的准确率

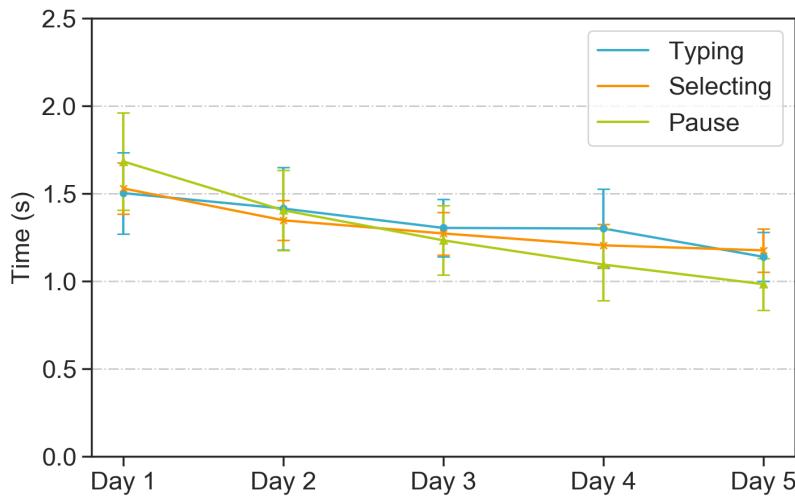
4.6.2.3 解码器准确率

如图4.16所示是通用化单词解码器和个性化单词解码器的准确率随实验天数而变化的柱状图。准确率指标包括 Top-1、Top-3 和 Top-5 准确率，其中 Top-N 准确率指的是被试所需单词出现在候选词列表前 N 位的概率。混合方差分析显示模型对 Top-1、Top-3 和 Top-5 准确率都没有显著性影响。实验天数对通用化模型下的单词解码准确率没有显著性影响，但对个性化模型下的 Top-1 准确率存在显著性影响 ($F_{4,28} = 3.45, p < .05$)。结果说明，个性化单词解码器会随着用户的使用而自我改进。混合方差分析显示，从第三天开始，个性化模型的解码准确率就存在超过通用化模型的趋势 ($F_{1,14} = 3.22, p = 0.09$)。

4.6.2.4 时间构成

为了深入探讨本技术的用户打字效率问题，实验者将被试的文本输入耗时拆分成三个构成部分：键入时间、选词时间和停顿时间。键入时间指的是被试点击单词各个字母所用的时间，是从被试点击单词首字母到点击单词尾字母之间的时间。选词时间是被试从候选词列表中选中所需单词的时间。停顿时间是从完成一个单词的选词到输入下一个单词首字母所需的时间。上述时间指的都是每输入一个英文单词所消耗的时间。

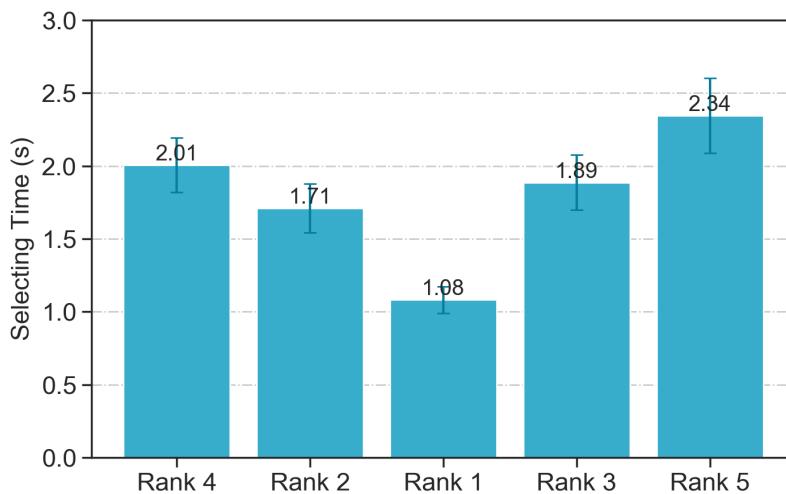
如图4.17所示是文本输入时间构成随着实验天数变化的曲线。方差分析显示，实验天数对键入时间 ($F_{4,52} = 8.01, p < .001$)、选词时间 ($F_{4,52} = 24.80, p < .001$)



图中展示了被试每输入一个英文单词时各个动作所消耗的时间，误差条表示 95% 置信区间。

图 4.17 文本输入时间构成

和停留时间 ($F_{4,52} = 18.86, p < .001$) 都有显著性影响。模型对上述时间构成都没有显著性影响。结果说明，被试在不断打字的过程中学会了更快速地键入字母和选中单词，同时不牺牲选词的准确率。



图中展示了被试在选中候选词列表中不同单词时的选词时间，误差条表示 95% 置信区间。

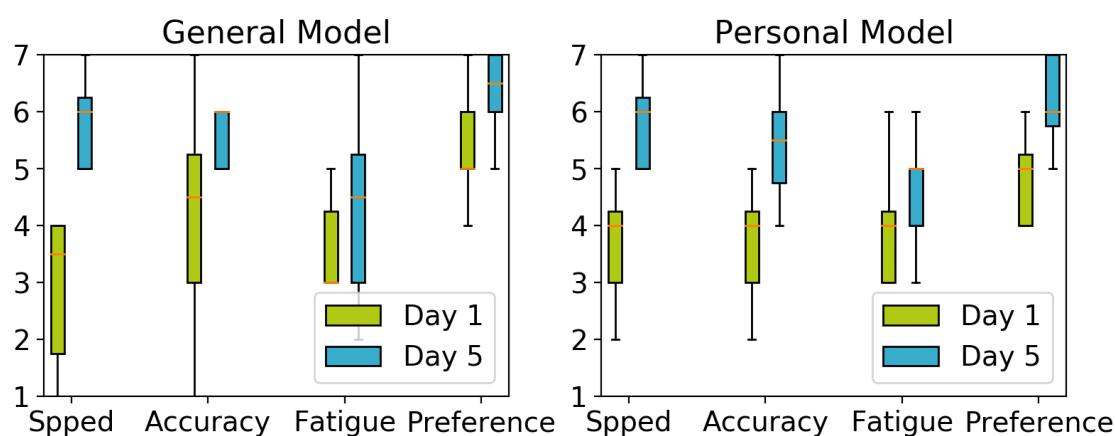
图 4.18 选词时间

图4.18进一步探讨了被试所需单词出现在候选词列表不同排位上时被试的平均选词时间。图中从左到右柱子和候选词列表中从左到右单词的顺序相对应，分别对应单词解码排名为第 4、第 2、第 1、第 3、第 5 的单词。从图中可以看出，

被试选择排名第一的单词所需的时间是最短的，这是因为在选词开始时，指针已经在该候选词上了。对于排名为2到5名的候选词，实验者发现有两个因素影响了被试的选词时间。第一个影响因素是该候选词距离中央候选词的距离 ($F_{1,15} = 69.47, p < .001$)。第二个影响因素是该候选词位于中央候选词的左侧还是右侧 ($F_{1,15} = 39.76, p < .001$)，被试选择左侧候选词比选择右侧候选词更快，这是人的手腕在左右旋转的限制上存在不对称性导致的^[155]。以上结果说明，在智能打字指环候选词区域的交互设计中，将排名第1的候选词放在中央，将排第2、第4名候选词放在左侧，第3、第5名候选词放在右侧，是十分科学的。

4.6.2.5 主观评分和反馈

在第一天和第五天的实验结束后，被试都通过一张七级李克特量表对使用本技术时的主观速度、主观准确率、疲劳程度和偏好程度打分（1分-最差；7分最好）。在第一天中，各项主观评价的分数就已经达到了可接受的程度，而且这些分数在第五天的时候得到了提高。对于通用化模型，Wilcoxon Signed-Rank 测试表明，主观打字速度 ($Z = -2.56, p < .05$) 和偏好程度 ($Z = -2.07, p < .05$) 在五天的使用过程中得到了改进。对于个性化模型，主观打字速度 ($Z = -2.21, p < .05$)、主观打字准确率 ($Z = -2.41, p < .05$) 和偏好程度 ($Z = -2.41, p < .05$) 在五天的使用过程中得到了改进。模型对所有主观评分都没有显著性影响。



图中展示了通用化模型和个性化模型下，实验第一天和第五天中，各项主观评价指标的分数。

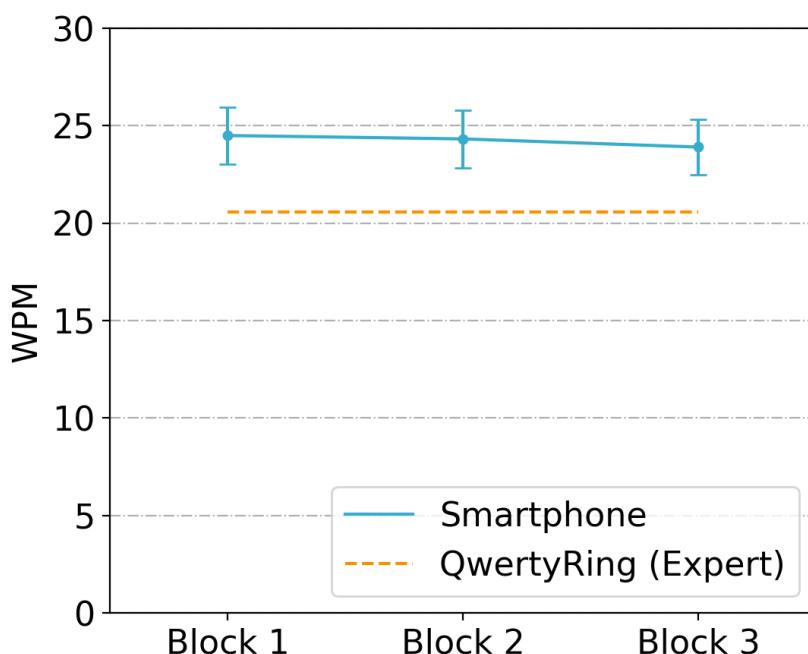
图 4.19 主观评分（1 分最差，7 分最好）

有两名被试在第一天的实验当中汇报他们在打字 20 分钟后感到疲劳，从第二天开始他们的疲劳感有所消退。实验观测发现，这两名被试心中的想象键盘布局太大了，这导致了疲劳的问题。当他们逐渐意识到想象键盘实际上可以小一些，并不影响打字的准确率时，他们感受到的疲劳程度就降低了。打字速度最快的被试第

一天的打字速度就达到 19.37 WPM，第五天时达到 29.12 WPM，他从第一天开始就严格遵循了实验者的指导。打字速度最慢的被试第一天打字速度为 7.54 WPM，第五天时为 16.67 WPM。这名被试犯了平移手腕的错误，这是智能打字指环使用方法中禁止的，会导致惯性传感指环很难获取表征用户点击位置的有效信息。然而在接下来的实验中，这名被试逐渐克服了困难，并且达到了一个可以接受的打字速度。

4.6.3 与手机打字速度进行比较

实验者组织了一场额外的用户实验来评测智能手机上食指打字的速度，实验的目的是比较智能打字指环和常用文本输入技术。实验者选择智能手机上的食指打字为基线技术，这是因为（1）智能手机上的文本输入是常用的，为被试所熟悉；（2）因为智能打字指环中被试使用食指输入文本，本实验也采用食指打字的方案。实验三中的 16 名被试参与了本实验，实验任务是分三段实验誊写 30 句话，每两段实验之间休息 1 分钟的时间。实验前，被试有十分钟的热身时间，正式实验大约持续了 20 分钟。



蓝线展示了被试在智能手机上食指打字的速度，误差条表示 95% 置信区间。橙色虚线是 QwertyRing 的专家打字速度，平均为蓝线的 86.48%。

图 4.20 比较智能打字指环与智能手指食指打字

图4.20展示了智能手机食指打字速度（蓝线）与智能打字指环专家打字速度的比较（橙线）。被试在三段实验中平均打字速度达到了 23.81 WPM。方差分析表面实验段数对打字速度没有显著影响 ($F_{2,30} = 0.20, p = .82$)，也就是说，被试们

已经达到了其手机食指打字速度的上线。智能打字指环的打字效率是这一结果的 86.48%，这表明智能打字指环提供了一种可穿戴的文本输入方法，而不会极大地牺牲打字速度。

4.6.4 对评测结果的讨论

4.6.4.1 与其它文本输入技术比较

智能打字指环的评测结果可以总结为：被试通过十分钟的热身练习，可以达到每分钟输入 13.75 个英文单词的打字速度，未纠正错误率为 1.5%；经过五天的练习后，被试可以达到每分钟输入 20.59 个英文单词的打字速度，未纠正错误率为 1.3%。这一表现明显优于现有的基于智能指环的文本输入技术。表格4.2的上半部分将智能打字指环与现有智能指环文本输入相关论文进行了比较。表格中的新手指标（Novice）是各论文中第一时段实验的评测结果，专家指标（Expert）是最后一个时段实验的评测结果。其中，智能打字指环和 *RotoSwype* 都进行了为期五天的用户实验，*ThumbText* 和 *TipText* 进行了至少 40 分钟的用户实验，从表格中可以看出，智能打字指环在新手打字速度和专家打字速度上都有很大的优势。表格4.2的下半部分展示了智能手机和智能手表上人们的平均打字速度，智能打字指环的表现接近智能手表上的打字速度，是智能手机上打字速度的三分之二。

	Novice	Expert	Novice	Expert
	WPM	WPM	UER%	UER%
智能打字指环 (通用化)	13.75	20.35	1.03	0.95
智能打字指环 (个性化)	13.74	20.83	1.90	1.63
RotoSwype (指环) ^[95]	9.2	14.8	2.4	0.5
ThumbText (指环) ^[125]	5.5	11.4	13.3	9.1
TypingRing (指环) ^[156]	NA	8.2	NA	NA
TipText (指环) ^[126]	10.5	13.3	0.3	0.3
Smartphone ^[157]	30.3		2.4	
Smartwatch ^[147]	22.0		1.5	

表 4.2 表格展示了智能打字指环与相关技术的对比。请注意，这一对比仅供参考，由于不同文献中的实验方法和实验条件不同，此对比不是完全公平的。

4.6.4.2 学习效应

智能打字指环用户需要将手腕贴在交互表面上，通过转动手腕来打字，用户需要适当的练习才能遵循此要求。在实验三的正式实验之前，有一个十分钟的热

身阶段，实验者向被试介绍了智能打字指环的概念与用法，然后被试在实验者的指导下进行练习，并在十分钟的热身之后，达到了 13.74 WPM 的打字速度。这说明，智能打字指环是容易上手的。在为期五天的联系之后，被试达到了 20.59 WPM 的打字速度，而且学习曲线并没有收敛。这说明，若想使用智能打字指环达到专家级打字速度，需要较长的学习过程，但是专家打字速度非常快。

4.6.4.3 通用化模型，还是个性化模型？

实验三在各个方面对比了智能打字指环的通用化模型和个性化模型，无论在打字速度、打字准确率、单词解码器准确率和主观反馈上，两个模型都不存在显著性差异。原因有两点：首先，通用化模型已经能很好地预测用户所需单词了，因此个性化模型的改进空间有限；第二，单词解码器的改进对打字速度的提高而言不是很关键。实验三中对打字时间构成的分析已经说明，被试打字变快的主要原因是他们键入字母和选词的速度更快了，而单词解码器的改进对打字速度的帮助不大。尽管如此，在通用化模型和个性化模型的选择上，实验者仍然推荐个性化模型。这是因为，实验三已经发现，从第三天开始，个性化候选词解码器呈现出由于通用候选词解码器的趋势 ($p = .09$)；被试的主观反馈也显示，他们感受到了个性化模型对单词解码准确率的改进。

4.6.4.4 若指环佩戴在食指第一指骨会怎么样？

在智能打字指环的交互设计中，实验者最终确定应当将指环佩戴在食指的第二指骨上。然而从社会接受程度的角度来说，各手指的第一指骨是更好的选择^[5]，那么如果将指环佩戴在食指第一指骨上，智能打字指环的表现会怎么样呢？不幸的是，若指环佩戴在食指第一指骨上，智能打字指环将无法支持文本输入。具体来说，这一交互设计让指环无法区分想象键盘的第二行和第三行字母，因为这两种情况下惯性传感器的信号非常类似。实验者邀请实验三中四位专家用户（打字速度前四名）进行尝试，将指环佩戴在第一指骨上打字（单词解码器也针对该设置更新），他们甚至无法完整输入一句话，因为他们想要的单词往往不在候选词列表中。

当实验者将词库大小减少到 100 个单词时，四个参与者可以输入单词。在这种情况下，贝叶斯解码器可以找到所需的词，因此，在某些使用场景中，将智能打字指环佩戴在食指第一指骨上也是可行的：用户可以通过使用智能打字指环输入语义命令来触发智能手机上的快捷方式，如“播放音乐”。

4.6.4.5 智能打字指环在其它无源表面上的表现

智能打字指环单词解码器的准确率是本技术能否应用在一个无源表面上的关键，它受到两个因素的影响：(1) 触摸事件识别的准确率；(2) 触摸事件发生时手指运动信号是否能表征用户所想字母。实验者认为，智能打字指环支持任何类似桌面的无源表面上的输入文本，类似桌面的表面指的是坚硬的、平坦的、宽敞的表面：首先，实验二已经证明，触摸事件识别技术在类似桌面的交互表面（如木质和塑料桌子）上表现良好；第二，当用户在平坦的表面上打字时，其触摸时刻的手指运动信号应该与实验环境相似，而本技术在实验环境下的变现良好。

然后需要承认的是，智能打字指环无法支持软性表面的文本输入，以坐姿大腿上侧为例，大腿上侧不够平坦和宽敞，这会严重影响惯性传感指环的运动信号。实验者邀请了实验三的四名专家用户（打字速度前四名）参与非正式实验，评测他们使用智能打字指环在木质桌面、塑料桌面和大腿上侧的打字速度。四名被试在木质桌面上誊写 10 句话的速度是 24.68 WPM ($SD=4.13$)，在塑料桌面上打字速度为 23.96 WPM ($SD=2.42$)。然而，被试们无法在大腿上侧完成所有的文本输入任务，因为一些想要的单词不在候选列表中。因此，大腿上侧基于惯性传感指环的触摸识别技术是可行的，但是基于指环的大腿上侧文本输入技术是不可行的。

4.6.4.6 电量问题

本章所介绍的内容默认以 1000 赫兹的惯性传感器信号作为基础，采用这一高频率信号的目的在于评测智能打字指环在最佳情况下的表现。然而，以 1000 赫兹的频率将惯性传感器数据发送到具有计算能力的机器上可能会很快地消耗智能指环的电量。有几种方法可以缓解电量问题：(1) 首先，智能打字指环也可以在 200 赫兹的惯性传感信号下运行。实验二结果显示，触摸手势识别算法在 200 赫兹下运行良好。对于文本输入法单词解码器而言，200 赫兹也是足够的，因为解码器只用到了传感器的俯仰角和偏航角等低频信息。(2) 第二，商用的智能指环有开关，用户可以在不使用智能指环时关闭电源。在更理想的情况下，未来的研究者可以为智能指环开发一个低功耗模式，在用户可能进行文本输入时才切换至高功耗模式。

Oura Ring (<https://ouraring.com/>) 是一个智能指环的例子，它比智能打字指环包含更多的传感器，包括红外接近光传感器、惯性传感器、体温传感器、蓝牙和电池。Oura Ring 的电池重 4 到 6 克，宽 7.9 毫米，厚 2.55 毫米，一次充电可使用一周。相比之下智能打字指环所需的传感器更少，如果实现商业化，其电量问题也将得到解决。

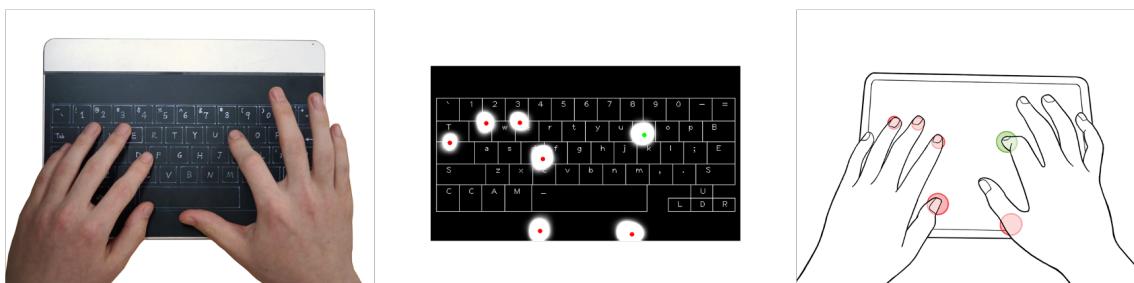
4.7 本章小结

文本输入是重要的人机交互任务，然而，新型显示设备如增强现实头盔、智能电视等等，都面临文本输入效率不高的问题。本章提出了基于惯性传感指环的无源文本输入技术，该文本输入技术与显示设备解耦，用户只需佩戴智能指环，即可在任意无源表面上高效打字，这一特性使得该技术适用于各类显示设备，具有很强的普适性。利用触摸运动模型，惯性传感指环能够有效识别触摸交互中的触摸事件和抬起事件，准确率分别为 99.30% 和 99.50%，还能准确分类多种触摸手势，如长按、滑动、左滑和右滑，准确率均高于 99%。实验者通过用户实验分析了被试使用惯性传感指环的打字行为，并据此设计了文本输入的单词解码器。在评测实验中，被试的上手打字速度为 13.75 WPM，经过五天的练习后打字速度达到 20.59 WPM。结果表明，基于惯性传感指环的无源表面文本输入技术易于学习，且打字效率远超其它智能指环文本输入技术，在增强现实、虚拟现实、智能电视等应用场景都具有实用价值。

第5章 平板电脑上的触摸意图推理键盘

5.1 引言

越来越多的人在平板电脑上使用文本输入，来完成诸如上网冲浪、写电子邮件和数字办公^[158]的交互任务。平板电脑上的文本输入法是一种基于触摸交互的软件键盘技术，这种软键盘在疲劳问题^[159]、视觉切换问题^[98,160-161]和打字速度^[152,162]上都与物理键盘存在巨大的差距。人在物理键盘上打字时可以将手指休息在键帽上，但在平板电脑触摸屏上打字时却不能休息手指，因为手指接触触摸屏会导致误触。手指的休息行为对物理键盘文本输入而言是至关重要的：首先，它可以缓解疲劳问题^[163]；第二，人能够通过键帽的触觉反馈来定位自己的手指，从而实现盲打^[164-167]。盲打指的是利用触觉（而非视觉）来定位手指在键盘上的位置，这可以大大提升文本输入速度。为了缩小触摸屏软件键盘和物理键盘之间的差距，本章将介绍 TypeBoard，一种面向平板电脑文本输入的防误触技术。如图5.1所示，绿色点是用户真正的打字触摸，红色点是误触。TypeBoard 能够过滤掉 99% 以上的误触，不仅能够过滤诸如小鱼际触碰这种常见的误触，还允许用户在打字的时候将非交互手指休息在触摸屏上，而不会引发误触。在强力防误触技术的支持下，未来的研究者可以通过可变形屏幕^[168]或可变表面纹理技术^[169-171]在触摸屏上提供触觉反馈，使用户在平板电脑上也能盲打。



TypeBoard 是面向平板电脑文本输入的防误触技术，能够过滤文本输入中的误触（红色点），只对用户的打字触摸事件（绿色点）作出响应。

图 5.1 TypeBoard 原理介绍

本文并不是第一份平板电脑文本输入的防误触工作。在 2013，Kim 等人就提出了 TapBoard^[163]，这是一款从规则上自定义了有意和无意触摸的平板电脑文本输入技术，该技术将“敲击”视为有意的触摸交互，而将其它触摸行为都视为误触。“敲击”指的是触摸时长小于 450 毫秒且触摸期间手指位移小于 15 毫米的点击，在该规则的约束下，用户需要主动适应键盘才能正常打字，这导致了不自然的

打字行为。而且，该技术必须在人的手指抬起时才能确定触摸的意图，并输入字母，相比之下，大多数文本输入技术都是在人的手指按下时就键入字母了。即便有以上种种约束，TapBoard 的准确率也只有 97%。本章所介绍的 TypeBoard 有三条策略来克服先前工作的不足：首先，为了确保交互的自然性，本章从人的交互意图出发，将误触定义为不表达交互意图的触摸。第二，本章将通过用户实验来了解用户的打字行为，并据此设计更好的面向平板电脑的防误触技术。第三，本章将利用触摸运动模型，寻找能够有效区分有意触摸和无意触碰的特征。

上面提到，用户行为规律对技术的改进有帮助，但是反过来，技术也会改变用户行为。举例来说，用户不会在打字的时候将手指休息在触摸屏上，除非触摸屏拥有很强的防误触能力。在设计 TypeBoard 的过程中，研究者需要观察用户的打字行为，然而，最有价值的用户行为数据正是用户在 TypeBoard 上的打字行为数据。这就面临一个循环论证的问题：研究者如何在研发出 TypeBoard 之前观察并分析用户在 TypeBoard 上打字的行为呢？为了解决上述问题，本章采用迭代的实验方法来研该防误触技术，并先提出三个研究问题来理清这份工作的思路：**(RQ1)** 用户在 TypeBoard 上的打字行为有何规律；**(RQ2)** 如何基于用户打字行为设计 TypeBoard 防误触技术；**(RQ3)** 用户使用 TypeBoard 打字的效率和体验如何？实验者组织了三个用户实验来回答以上研究问题。考虑到 RQ1 和 RQ2 是循环论证的，实验者采用迭代的实验方法来逼近这两个问题的答案：

1. **实验一：**用于回答 *RQ1* 和 *RQ2* 的第一轮迭代。在实验一中，实验者收集了被试在没有反馈的触摸板上的打字数据。被试不能真正输入单词，而是想象可以输入单词，并且想象该触摸板能够防止所有误触。在这个实验中，实验者共观察到了 11 类误触行为，例如多指休息、小鱼际误触和边缘误触，等等。基于实验一的数据，实验者研发了初版 TypeBoard。
2. **研究二：**用于回答 *RQ1* 和 *RQ2* 的第二轮迭代。在实验二中，实验者收集了被试在初版 TypeBoard 上的打字数据。结果显示，实验一中观察到的 11 类误触行为已经是误触类型的全集了，在实验二中未发现新的误触类型。然而，实验二的用户行为在细节上与实验一有一些差异。例如，在实验二中，有意触摸的平均压力减轻了 33.78%，这是因为在有反馈的情况下，被试们可以根据反馈用更轻的力气打字，以避免疲劳。基于实验二的数据，实验者研发了终版 TypeBoard，防误触准确率达到了 98.88%，识别延迟为 100 毫秒。
3. **研究三：**对 TypeBoard 进行评测，回答 *RQ3*。在实验三中，我们评测了被试在三种设置下的打字效率和用户体验，包括普通平板电脑软键盘、TypeBoard 和 TypeBoard+（有按键轮廓反馈的 TypeBoard）。结果表明，TypeBoard 通过

缓解疲劳问题、减轻认知负担、降低打字错误率大大提升了平板电脑上文本输入的效率和可以用性。TypeBoard+ 则是进一步的改进，让用户能够在平板电脑上盲打。

本章工作有三点主要贡献：首先，本章提出了面向平板电脑文本输入的防误触技术，简称 TpyeBoard，能以 98.88% 的准确率、100 毫秒的延迟过滤触摸交互中的误触；第二，本章整理了平板电脑文本输入的用户打字行为；第三，本章通过用户实验证明，TypeBoard 提升了平板电脑文本输入的交互体验，具体体现在打字效率、打字错误率和用户主观反馈三方面。

5.2 相关工作

5.2.1 触摸交互中的防误触技术

触摸交互是智能手机、平板电脑和智能大桌面等触摸屏设备的主要输入信道。并非所有触摸屏上的手指接触都应该触发数字响应，那些对用户的交互目标没有帮助的触摸被称为误触^[172-174]。如果误触触发了不必要的数字响应，就会影响正常交互的效率和自然性^[172,175]。不必要的数字响应会破坏用户的交互流程，用户需要花费额外的时间和操作来消除该响应的影响。随着时间的推移，用户会倾向于在屏幕屏设备上小心地进行触摸交互，以避免误触。例如，在普通的平板电脑上，用户打字时会将非交互手指悬停在触摸屏上，以避免误触，这种将手悬空的做法会导致疲劳问题^[176]。

尽管误触对交互有负面影响，但是误触在触摸屏交互中是不可避免的。例如，在人们日常使用智能手机的过程中，人手的大鱼际就会不断地触碰屏幕的边缘^[177]。幸运的是，各种各样的软件技术有机会过滤误触。在相关文献中，已经有广泛的工作致力于研究防误触算法，本节将通过两种分类方法来介绍相关工作：（1）应用场景，和（2）传感器类型。

5.2.1.1 不同应用场景下的防误触技术

在不同的应用场景下，误触的定义也会有所不同。若不对应用场景作出限制，误触指的是不表达交互意图的触摸^[172]；而在特定的交互任务中，误触的定义变得更加明确，例如，文本输入任务中的误触指的是不用于输入单词的触摸。

少量文献^[172,174]和大量专利^[178-182]在不分应用场景的情况下识别误触。Metero 等人整理了智能手机上的误触行为，并提出了智能手机上的防误触算法^[174]，该方法能过滤 79.6% 的误触。Xu 等人利用眼动信号、头动信号和触摸屏触点信息识别触摸屏大桌面上的误触，准确率为 91.3%。上述方法的特点是适用于不同的应

用场景，但是识别准确率较低。

在相关文献中，更多的实验致力于研究特定应用场景下的防误触问题。TapBoard^[163]和 TapBoard2^[183]在触摸屏文本输入任务下讨论误触问题。TapBoard 将敲击手势识别为有意触摸，将其它所有触摸都识别成误触，其中，敲击手势指的是触摸时长小于 450 毫秒，触摸期间手指位移小于 15 毫米的触摸。用户需要主动适应技术中所规定额阈值来进行打字，这导致了不自然的用户行为。即便如此，TapBoard 的误触识别率也只有 97%。TapBoard2 进一步区分了打字触摸和用于目标选择的触摸，准确率为 95%。在基于触摸屏和电子笔的应用场景下，误触是最被诟病的问题之一^[175]，例如，在使用防误触能力有限的产品时，用户必须保持别扭的握笔姿势，以防止误触。有几份工作致力于研究电子笔场景下的防误触问题^[184-186]。Schwarz 等人利用触摸的时空特征来训练误触过滤器，其准确率是同类型研究中最高的，能够将误触和有意交互的比例降低到 0.016，同时正确识别 98% 的有意交互。总体上来说，特定应用场景下的防误触技术准确率更高。

5.2.1.2 不同传感技术下的防误触技术

大量相关工作致力于研究触摸屏设备上的防误触方法，相关设备包括智能手机^[173-174,177,187]、平板电脑^[163,179,183,185-186]和触摸屏大桌面^[172]。大多数技术都仅利用触摸屏的触点信息^[163,174,179,183,186]和电容屏图像^[177,185]。Metero 等人利用触点信息识别智能手机触摸屏上的误触^[174]。他们提出了基于触摸时长、触摸位置分布和轨迹特征的误触过滤标准，成功过滤了 79.6% 的误触，对有意触摸的未识别率仅为 0.8%。Schwarz 等人提出了平板电脑上区分有意的电子笔输入和手掌误触的方法^[186]，他们从触摸点中提取特征并使用决策森林训练机器学习模型。PalmTouch^[177]可以区分智能手机上的有意触摸和手掌误触，它利用触摸屏的电容图像作为输入，使用卷积神经网络（CNN）训练模型，在实测场景中达到了 99.53% 的误触识别率。

另有一些工作通过增加传感能力来提高防误触能力^[172-173,187-188]。GestureOn^[173]能在智能手机处于待机模式的情况下区分有意的手势输入和误触，使得用户可以在打开屏幕之前通过手势触发快捷方式。GestureOn 利用了绝大多数智能手机中的内置传感器，包括红外接近光传感器、惯性传感器、屏幕上的压力信号等等，通过多信号融合技术，GestureOn 误触识别达到了 98.2% 的精确率和 97.6% 的召回率。Xu 等人利用眼动信号、头动信号和触摸屏触点信息来识别触摸屏大桌面上的误触，结果表明，眼动信号和头动信号使得误触识别的准确率提升了 4.3%，达到 91.3%。总体来说，更多的传感信号的融合总是有利于提高防误触能力。

5.2.1.3 小结

本小节提出两个观点：首先，一般来说，防误触方法总是在场景受到限制的情况下效果更好。在特定场景下，有的防误触技术的准确率相当高，以至于让用户改变自身交互行为并从中收获更自然的交互体验^[163,175,183]。第二，多种传感信道的融合通常可以提高防误触能力。多项研究表明，额外的输入信道（例如压力信号^[173]、眼动和头动信号^[172]）有助于提高误触的识别率。在本章所述的面向平板电脑文本输入的防误触技术中，研究者也使用了压敏触摸屏，并将应用场景限定在文本输入交互任务中，在这两条前提下本技术应该能具备很强的防误触能力。

5.2.2 触摸交互防误触技术的作用

防误触技术的直接好处是降低误触对交互产生的负面影响，除此之外，本小节将防误触技术的其它用处总结如下。

5.2.2.1 弥补触摸屏文本输入之于物理键盘的不足

越来越多的人需要用到平板电脑的软键盘文本输入技术^[158]，然而，平板电脑的软键盘在打字效率^[152,162,189]和用户体验^[98,159-161]上无法与物理键盘相提并论。在物理键盘上，用户可以将手指休息在键帽上，这是物理键盘的重要特性：首先，这可以缓解长时间打字的疲劳问题^[163]；第二，用户通过触摸键帽的轮廓来定位自己的手指位置，从而通过盲打提高打字效率^[164-167]。在平板电脑的软键盘上，若能实现一种强大的防误触技术，使得用户能在打字时将非交互手指休息在触摸屏上，将能显著缓解疲劳问题。更进一步地，在手指休息不会触发不必要的误触的前提下，研究者可以通过可变形触摸屏^[168]或可变触摸屏纹理技术^[169-171]来模拟物理键盘上键帽的触觉反馈，从而使得用户在平板电脑上也能盲打。如此一来，触摸屏软键盘和物理键盘之间的差距就会缩小。

5.2.2.2 将误触输入作为有用的信息

大部分与误触相关的工作的目标都是去除误触，然而，一部分试图将误触应用于情景感知^[190]和有意交互^[177,183,191]。手掌在平板电脑上的按压通常被视为误触，但 Zhang 等人手掌按压作为平板电脑上一种新的触发命令的方式^[190]，增强了平板电脑上的电子笔交互。例如，当用户在用电子笔绘画时将手掌按压在屏幕上，在手掌周围就会出现一圈绘画工具，方便用户的使用。前臂本是触摸屏大桌面的典型误触部位，但 Koura 等人将前臂的触碰应用于触摸屏大桌面的菜单操作中^[192]。Matulic 等人将各种手掌与触摸屏的接触姿态用于丰富触摸屏大桌面的交

互，该系统以 91% 的准确度分类七种不同的接触姿态，可用于触发和控制菜单和小工具部件。PalmTouch 可以区分手指触摸和手掌触摸，并将手掌触摸定义为一种快捷触发方式^[177]。TapBoard2 可以区分打字和普通触摸行为^[193]，从而统一了软键盘和触摸板的空间，省去了用户在不同设备之间来回切换的成本。

5.3 初版防误触键盘

5.3.1 实验一：收集无反馈键盘上的打字数据

实验一采集了用户在一块压敏触摸板上的打字数据，这一实验是探索用户在 TypeBoard 上打字行为的第一轮迭代，研究用户打字行为有助于防误触技术的设计，基于本实验的数据本节将介绍初版防误触键盘技术。由于（不正确）的反馈可能会影响用户的行为，在这个实验者，被试们在无反馈的条件下打字：被试们并非真的输入单词，而是想象他们的打字行为能够输入单词，同时，被试需要想象该触摸板能够完美地防止误触。

5.3.1.1 被试

实验者从校园中要求了 16 名被试参与实验，实验者的年龄从 19 岁到 26 岁不等，平均年龄 22.13 岁，标准差为 2.13，其中，有 8 名女性被试。所有被试都是右撇子，他们使用智能手机触摸屏文本输入的时间不少于两年，平均 7.50 年，标准差为 2.25。所有用户都有物理键盘打字经验，有 8 名被试有使用平板电脑打字的经验。

5.3.1.2 实验设备

如图5.2所示是实验设置，实验者将微软的 Surface 平板电脑和 Sensel 压敏触摸板^[79]拼在一起，以代替还未商业化的压敏触摸显示屏。压敏触摸板包含 185×105 个传感元件，间距为 1.25 毫米。每个传感单元可以感应到大约 30000 个级别的压力，范围从 5 克到 5 千克不等。压敏触摸板提供压力图像和触摸点信息，包括触摸点的位置、时间戳、面积、压力和拟合椭圆的参数。压敏触摸板的感应区域大小为 240 毫米 \times 138 毫米。实验者使用荧光笔在触摸板上手工绘制了二十六键键盘布局，以模拟触摸显示屏上的键盘布局。由于压敏触摸板的宽度（240 毫米）比典型笔记本电脑（15 英寸苹果电脑）上的键盘宽度（270 毫米）小。实验者在绘制键盘布局时移除了一些不常用的按键，例如方括号和分号，如此一来，二十六键键盘的布局就可以完整地绘制在压敏触摸板上了，而每个按键的大小与苹果电脑保持一致（17 毫米）。平板电脑是微软 Surface Pro6，装配 i7 英特尔酷睿处理器。实验所

用的程序以 50 赫兹运行。

5.3.1.3 实验设计和过程



实验者将一个平板电脑和一个压敏触摸板拼接在一起，以模拟带有压敏触摸屏的平板电脑。被试在无反馈的情况下想象在平板电脑上打字。

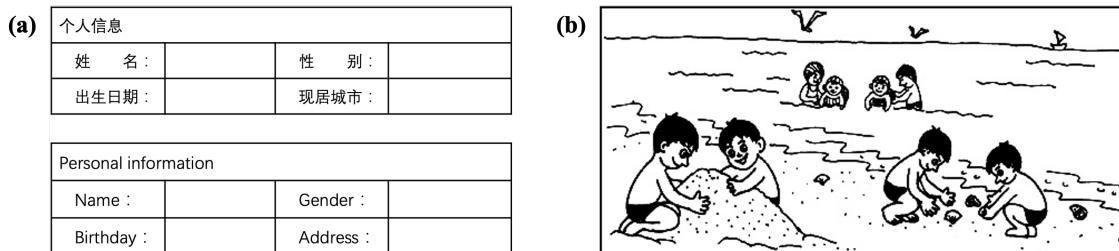
图 5.2 实验一的实验设置

实验任务是一份填写一份 Word 表格，该表格显示在平板电脑上。被试通过平板电脑上的触摸交互来选择表格中的单元格，通过在 Sensei 压敏触摸板上连续触摸点击来想象打字的情景。被试需要按照实验任务做出打字的动作，只是平板电脑不会对打字动作作出反馈。实验表格包含四个任务：(1) 填写个人信息；(2) 简短的问题；(3) 模拟开卷考试；(4) 看图写话。实验语言为中文。实验通过拉丁方来平衡以上四个任务的出场顺序。与典型的文本输入相关文献^[151,194-195]不同，本实验未包含誊写任务，这是因为实验者通过预实验发现，在誊写任务中被试很少会将手指休息在触摸屏上，因此誊写任务将不利于误触数据的收集。

1. **填写个人信息：**实验表格中有十个空的单元格属于填写个人信息任务，填写的内容包括姓名、性别等等。如图5.3(a) 所示是个人信息任务的图例和英文翻译。这一项实验任务代表那些需要频繁切换文本输入和光标控制的交互任务。为了保护被试的隐私，被试们可以在表格中填入虚构的个人信息，前提是他们需要记得自己填了什么，因为记得所填写的内容对后续的人工标注过程至关重要。
2. **简单的问题：**实验表格中包含十个简短的问题，例如“你最喜爱的颜色是？”、“你最好的朋友是？”。同样地，为保护被试的隐私，被试可以在表格中给出虚构的回答。
3. **模拟开卷考试：**实验表格中的模拟开卷考试部分包含五个很难的问题，例如，“元素周期表中第 50 号元素是什么？”。被试很有可能不知道问题的答案，

因此他们需要使用搜索引擎搜索答案。在本实验中，被试只是想象打字，他们边做出打字的动作边将打字内容说出来，实验者代替他们在浏览器中搜索相关内容。这一项实验任务代表需要浏览网页的一类交互任务。

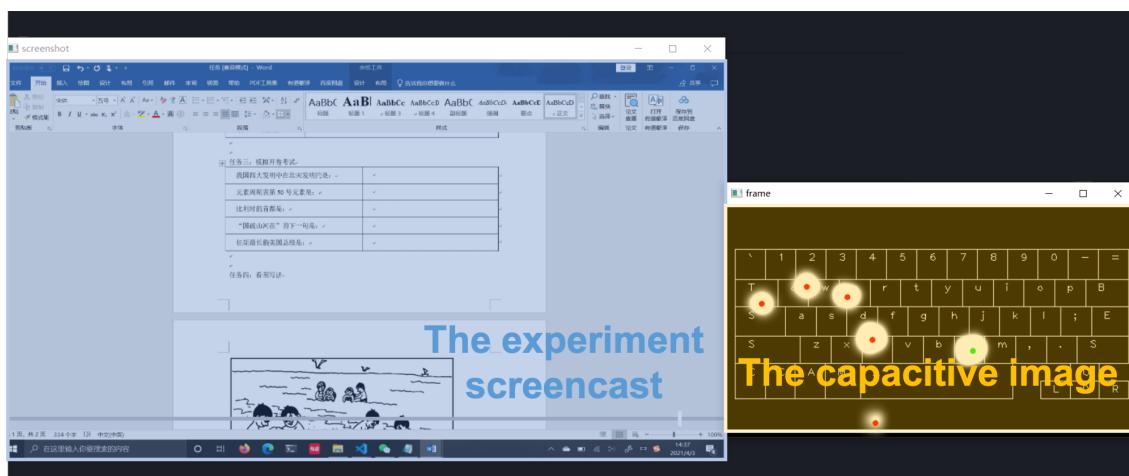
- 看图写话：如图5.3(b)所示，被试通过五句话来描述图中的场景。被试边做出打字的动作边将打字内容说出来，实验者代替他们输入文字。这一项实验任务代表自由写作这一类交互任务。



如图所示是实验表格中的任务。左侧是填写个人信息部分的内容，右侧是看图写话部分的内容。

图 5.3 实验任务图示

在正式实验开始之前，被试有五分钟的练习时间熟悉实验流程。在练习阶段，被试可以在触摸板上自由打字，打字过程没有反馈。在被试自由练习的过程中，实验者会提醒被试两点注意事项。首先，触摸板上的键盘没有任何反馈，被试不能真的输入文本，而只是做出打字动作，想象文本被成功输入。第二，被试需要想象该触摸板键盘能够过滤所有的误触，并据此改变自己的打字行为，例如，当被试在思考打字内容的时候可以将手指休息在触摸板上。



如图所示是人工标注过程中显示器中的内容。左侧是实验录屏，右侧是触摸板记录的触点数据的可视化，被试在右侧的可视化程序中给触点的有意性做人工标注。

图 5.4 标注程序图示

在每一项实验任务之后，被试都会给刚刚实验中触点的有意性做人工标注。如

图5.4所示，被试通过一个交互式程序给触点数据进行标注，程序既展示了实验过程中的录屏，也展示了触点数据的可视化图像，两个图像的时间轴是对齐的。可视化图像上的每个红色点表示一个触点，被试通过回忆来把其中的有意触摸标记成绿色。如果被试忘记了某个触点的有意性，可以将触点标记成蓝色，表示删除该数据。这个人工标注的准确性是可以保证的，因为被试可以通过实验录屏中的上下文信息来帮助回忆。平均而言，被试总共花10分钟的时间来完成实验表格中的内容，花45分钟的时间来给实验收集的触点数据做人工标注，被试在每两项实验任务之间都会休息5分钟的时间，因此实验的总时长约为70分钟。

5.3.2 基于打字行为设计初版防误触键盘

本实验收集的数据集包含12659个触点，不包括人工标注中被剔除的数据点。数据集包含67.5%的有意点击样本和32.5%的误触样本。基于该数据集，实验者分三个阶段研发了初版TypeBoard防误触技术，该技术基于支持向量机。如表格5.1所示是防误触技术的总览，展示了该技术的三个阶段中所涉及的机器学习特征和测试准确率。每一个特征组都是针对一种或者多种有意性容易混淆的情形而设计的，可以直接提升误触识别的准确率。本小节的剩余内容将分三个阶段详细介绍防误触技术的研发过程。

阶段	分类准则	特征	介绍	分类准确率
一	当前触点的信息	接触面积	手指与交互表面的接触面积（单位：平方厘米）	96.86% (SD=4.17%)
		拟合椭圆率	手指与交互表面的接触区域拟合成椭圆	
		触点位移	手指从按下和抬起的时间段内触点移动的距离	
		触点压力	手指按压交互表面的压力大小（单位：克）	
		触点压强	触点压力与接触面积之比	
二	当前触点的各种指标占所有触点的比例	压力占比	本触点压力占所有触点总压力的比例	98.59% (SD=1.46%)
		压强占比	本触点压强占所有触点总压强的比例	
		接触面积占比	本触点接触面积占所有触点接触面积的比例	
三	当前触点在触摸板上的位置	到触摸板边缘的距离	本触点到触摸板边缘的最小距离，边缘包括触摸板的下侧和两侧	99.07% (SD=0.71%)
		到触摸板角落的距离	本触点到触摸板角落的最小距离，角落包括触摸板的左下角和右下角	
三	时空相关性特征	与时域临近点之间的关系	本触点与五秒以内最近的五次触点之间的关系，包括（1）触摸时间戳之差；（2）平均距离；（3）压力之比；（4）接触面积之比；（5）压强之比。若五秒以内临近的触点不足五个，用零补足特征。	99.07% (SD=0.71%)
		与空间临近点之间的关系	本触点与空间上最近的五个触点之间的关系。若此刻触摸板上的触点不足五个，用零补足特征。	

表5.1 表格中展示了三个阶段中喂入支持向量机模型的特征组。除了“时空相关性特征”以外，其余特征组都需要提取100毫秒内的时域特征，包括最大值、最小值、平均值、峰度和偏度。

5.3.2.1 阶段一：用简单模型检查标注效果

由于一些被试误解了误触的概念，数据集中存在一些被错误标记的触摸点。实验者通过简单的特征提取来训练一个简单的模型，并使用该模型来负荷人工标注的结果。实验者对每次触摸事件后的连续五帧（100毫秒）进行采样，如果触摸持续时间短于五帧，则对整个手指接触交互表面的过程进行采样。如表格5.1（阶段一）所示d，实验者从样本中提取了如下特征：对于手指接触面积、拟合椭圆的椭圆率、位移、按压力度和压强，计算时域特征，包括五帧中最大值、最小值、平均值、偏度和峰度，然后将这些值拼接起来，得到一个25维的特征向量，喂入支持向量机中训练识别误触的二分类模型。这就是阶段一的防误触技术。

实验者用该模型来对数据集中每一个触点的有意性进行分类，发现其中一部分被试误解了误触的概念，例如，将打字过程中打错字母的情况也算成误触了。实验者通过电子邮件要求被试重新标注这些可疑的数据点。至此，实验者认为数据的标注是真正可靠的了，因为数据标注经过了被试、实验者和简单分类模型的三重校验。校验过后还剩下12624个数据点，包括68.38%的有意触摸样本和31.62%的误触样本。留一（被试）交叉验证表明，阶段一防误触技术的准确率为96.86%（SD=4.17%）。

5.3.2.2 阶段二：过滤多指休息行为导致的误触

实验数据分析发现，大部分误触来自被试的多指休息行为，这种行为导致的误触占误触总量的72.66%。在本实验中，实验者告知被试，应当想象该触摸板具有完美的防误触能力。在此背景下，被试的确会在打字过程中将手指休息在触摸板上，以避免疲劳。由于被试将手指休息在触摸板上的行为并不表达打字的意图，因此这些手指接触交互表面时带来的触点都属于误触，应该被防误触算法过滤。在这样的用户打字行为之下，过滤多指休息行为导致的误触问题就变得尤为关键。如表格5.1（阶段二）所示，实验者添加了一类触摸意图分类准则（占比），用于过滤多指休息行为导致的误触。在多指休息行为中，手指往往是陆续接触到交互表面的。在大部分情况下，在第一根手指触摸之后，后续的手指会在100毫秒以内接触交互表面：在86.05%的样本中，第二根手指会在100毫秒以内接触交互表面；在76.30%的样本中，第二、第三根手指都会在100毫秒以内接触交互表面。由于本模型的考虑触摸发生后五帧的数据，识别延迟为100毫秒，该模型有很大的机会来正确识别一次多指休息行为，从而将这些触摸点判定为误触。如表格所示，实验者添加了三个特征组来过滤多指休息行为导致的误触，这三个特征组分别对应触点的压力、压强和接触面积占此刻所有触点的相应属性之和的比例，对占比求

时域特征即得到特征组。留一（被试）交叉表明，阶段二防误触技术将误触的识别准确率提升至 98.59% ($SD=1.46\%$)。实验者用两种指标来衡量该防误触技术的错误率：(1) 假正率——将一次误触识别成有意触摸的概率；(2) 假负率——将一次有意触摸识别成误触的概率。本防误触技术的假正率为 3.16%，假负率为 0.59%。

5.3.2.3 阶段三：理解用户打字行为并深入改进技术

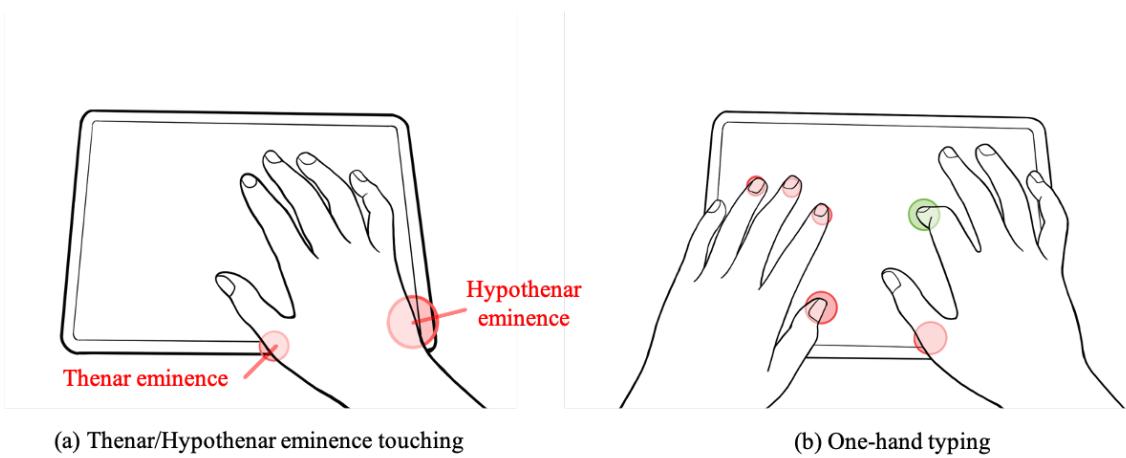
实验者通过分析阶段二防误触技术中预测失败的案例来理解用户打字行为中容易导致误触的部分，从而指引防误触技术的改进方向。如表格5.2所示，实验者将预测失败的案例分为 16 类，包括十种容易误触的情况，和六种容易被识别为误触的有意触摸。实验者对每一类预测失败的案例进行计数，讨论其原因，并给出可能的解决方案。对于表格中白色部分的触摸案例，实验者认为机器有能力正确识别其有意性，对于这些案例，实验者提出了新的特征组来提升模型的预测准确率。下面是一些例子：

1. **例一：小鱼际误触。**如图5.5 (a) 所示，小鱼际指的是控制小指运动的肌肉群，而大鱼际指的是在拇指根部上的肌肉群。大鱼际误触和小鱼际误触都是触摸交互中常见的误触类型。而在平板电脑十指打字这一交互任务下，小鱼际误触更是十分常见，这是因为在十指打字过程中小鱼际经常会以较大的力度触碰到触摸屏，这很容易引起误触。幸运的是，小鱼际误触总是出现在触摸板的左下角或者右下角，因此，“与触摸板四个角的距离”这一数值成为过滤小鱼际误触的有效特征。
2. **例二：连续触摸。**当一名被试连续触摸键盘上的同一按键（如删除键）时，从第二次触摸开始点击的力度会变得更轻 ($p<.005$)。实验数据分析表面，在对同一个按键的连续触摸中，第一次触摸的平均力度为 167.01 克 ($SD=106.76$)，而后续触摸的平均力度为 150.42 克 ($SD=112.70$)。由于后续触摸更轻，它们更有可能被误识别为误触，因此时域上临近触点的信息将有利于对触摸有意性的判断。
3. **例三：单手打字。**如图5.5 (b) 所示，有时候被试会用单手打字，另一只手休息在触摸屏上。在这种情况下，阶段二的防误触算法可能会认为此刻所有的触点都是多种休息行为导致的误触，但实际上正在打字的手指所触发的是有意触摸。由于打字手指和休息中手指之间距离往往比较远，因此空间上临近触点的信息将有利于对触摸有意性的判断。

然而，对于表格5.2中灰色部分的触摸案例，即使是人类（实验者）也无法在不观察实验录屏的情况下区分其有意性。实验者任务，这部分触摸案例被误识别是无法避免的。这里有一些例子：第一，极少数情况下被试会将一只手指重重地休

案例	介绍	N	有用的特征
假正（容易被识别成有意触摸的误触）			
小鱼际误触	如图5.5 (a) 所示	23	触点位置，与触摸板角落的距离
大鱼际误触	如图5.5 (a) 所示	2	触点位置，与触摸板边缘的距离
时域重复报点	一次触摸被系统汇报了两次	12	时域临近点的信息
空间重复报点	一次触摸被系统识别成了两个触点	3	空间临近点的信息
边缘误触	用户在抓握设备时常有的误触	7	触点位置，与触摸板边缘的距离
二指休息	两只手指休息在屏幕上	3	空间临近点的信息
轻触误触	输入单词时多出来的一次轻轻的点击	9	该触点是否比最近的触点更轻
滑动误触	滑动通常不表达打字意图	7	触点的位移
单指休息	两只手指休息在屏幕上	9	无解
重触误触	输入单词时多出来的一次较重的点击	15	无解
假负（容易被识别成误触的有意触摸）			
连续触摸	连续点击同一个按键时，后续点击变轻	13	时域临近点的信息
粘滞打字 ^[196]	在前一只手指抬起之前，下一只手指已按下	12	时域临近点的信息
单手打字	一只手打字，另一只手多指休息	6	空间临近点的信息
手掌触摸	一只手指打字的同时，手掌造成了多个误触点	5	空间临近点的信息
轻触触摸	很轻但有意的触摸	46	无解
小面积触摸	触摸面积很小但有意的触摸	6	无解

表 5.2 阶段二防误触算法的预测失败案例，N 表示失败案例计数。



左图展示了小鱼际误触和大鱼际误触，右图展示了一手打字、一手休息的情况。

图 5.5 有意性易混淆的案例

息在触摸板上，这一类误触会与有意触摸混淆；第二，在被试输入一个单词的过程中，有时候点击其中一个字母的力度极轻，这一类有意触摸会与误触混淆。因此，根据对表格中灰色部分触摸案例的计数，实验者认为，防误触技术在本数据集上存在一个准确率的理论上限，约为 99.40%。

根据表格5.2中汇总的有意性易混淆的触摸案例，实验者进一步添加了两条触摸意图分类准则，形成阶段三的防误触技术。第一条分类准则与触点的位置有关，包括触点距离边缘和角落的最短距离。实验者并未将触点位置与二十六键键盘布局的关系考虑进去，这是因为触摸屏上软键盘的布局是可以变化的，而实验者希望得到的是一个与键盘布局解耦的通用化防误触算法。第二条分类准则是连续触摸的时空相关性特征，包括本触点与时域临近点之间的关系，和本触点与空间临近点之间的关系。具体的特征描述在表格5.1（阶段三）中。阶段三防误触算法用到表格中的所有特征，拼接起来形成刚好 100 维的特征向量，喂入支持向量机中训练识别误触的二分类模型。留一（被试）交叉验证显示，误触识别的准确率为 99.07%，假正率为 1.77%，假负率为 0.54%。至此，实验者已经得到了一个高准确率的误触识别分类器。然而，正如本章在引言中所述，实验一所采集的用户打字行为数据并不能完全代表用户在强力防误触键盘上的打字行为，因此，实验者将阶段三的防误触技术命名为初版 TypeBoard。下一个实验将采集被试在初版 TypeBoard 上的打字行为数据，并使用新的数据来改进技术。

5.4 终版防误触键盘

5.4.1 实验二：收集初版防误触键盘上的打字数据

实验二收集了被试在初版防误触键盘上的打字行为数据，其中，初版防误触键盘是基于实验一数据的技术。实验二是探索用户在 TypeBoard 上打字行为第二轮迭代，研究该用户行为有助于防误触技术的改进，基于实验数据本节将介绍终版防误触键盘技术。

5.4.1.1 被试

实验者从大学校园中邀请了 16 名被试参与实验，被试年龄从 19 岁到 37 岁不等，平均年龄为 22.19 岁，标准差为 4.55，有 6 名女性被试。所有被试都是右撇子，没有参加过先前的用户实验。被试们都至少有一年的智能手机软键盘输入法使用经验，平均经验为 5.88 年，标准差为 2.36。10 名被试曾经使用过平板电脑上的文本输入法。

5.4.1.2 实验设计和过程

如图5.6所示，实验设备包括一台微软 Surface 平板电脑、一块 Sensel 压敏触摸板和一副耳机。实验共包含四项实验任务，任务与实验一相同，分别是在实验表格上填写个人信息、回答简单的问题、模拟开卷考试和看图写话。实验二与实验一的不同点是：（1）被试可以通过防误触键盘打字，而不是像实验一中一样只能在无反馈的键盘上做打字动作；（2）键盘具有一定的防误触能力，初版防误触键盘技术能够过滤大部分误触。在实验中，被触摸板判定为有意的触摸都会触发输入法响应，并有声音反馈，而被触摸板判定为误触的触摸不会引发任何响应。



实验二的设备包括一台平板电脑、一块压敏触摸板和一副耳机。

图 5.6 实验二的实验设置

在正式实验之前，被试有五分钟的时间通过自由打字来熟悉实验设置。由于所有被试都没有在强力防误触触摸屏键盘上打字的经验，实验者提醒被试留意，在打字过程中可以将非交互手指休息在键盘上。被试可以根据自己的个人喜好和打字状态选择休息或不休息手指。在每项实验任务之后，被试像实验一一样标注每个触点的有意性。平均而言，每名被试总共花 15 分钟在文本输入任务上，花一个小时在数据标注上。在每两项实验任务之间，被试休息五分钟的时间，以避免疲劳。整个用户实验用时 90 分钟，时间比实验一长，这是因为在实验二中被试需要真正输入文本，选词和纠错消耗了更多的时间。

5.4.2 基于打字行为设计终版防误触键盘

排除了标注过程中无法辨认有意性的触点后，实验数据共包含 13789 个触点，其中包含 71.01% 的有意触摸和 28.99% 的误触。实验一之后训练得的初版防误触技术在实验二的数据集下准确率为 98.05% ($SD=1.51\%$)，假正率为 1.35% ($SD=1.28\%$)，假负率为 2.20% ($SD=1.93\%$)。初版防误触技术在实验二中的准确率与在实验一中的准确率相比，下滑了一个百分点，这说明实验一采集的数据和实验二采集的数据之间存在差异。由于实验二的实验设置更接近真实场景，本小节有必要将两个实验中的数据差异讨论清楚。

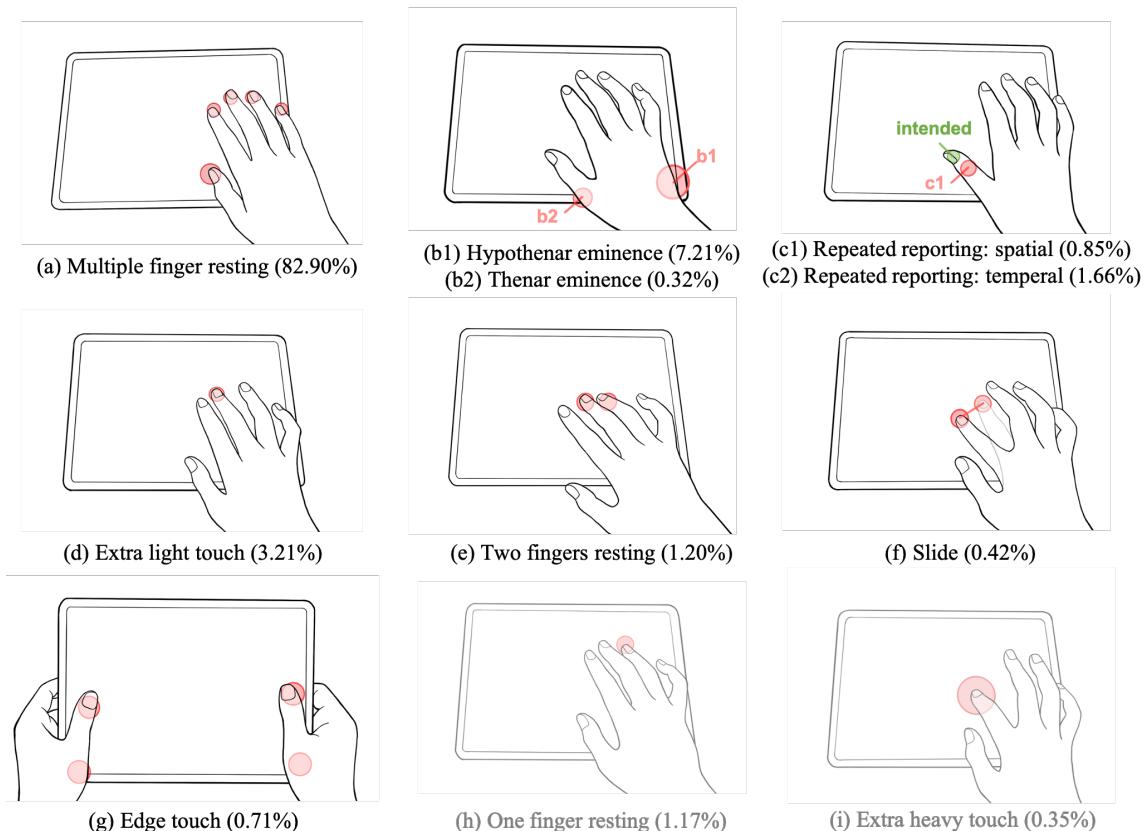
与实验一相比，实验者并未发现新的误触行为类型。然而，在用户行为细节上两个实验的数据存在差异。表格5.3中举起了一些例子。首先，实验二中被试的触摸力度比显著比在实验一中的更轻 ($t_{15} = -2.78, p < .01$)，在实验二中，被试是在有反馈的环境下打字的，当他们发现更轻的触摸依然能够成功触发输入法的响应后，便会逐渐减轻按键的力度，以节省体力。第二，在实验二的数据中，存在更多连续触摸的情况 ($t_{15} = 3.57, p < .005$)，即连续两次触摸之间的位移小于 0.5 个按键宽度、时间间隔小于 500 毫秒。这一现象主要是因为在是实验二中，被试需要连续点击退格键来删除错误的单词。第三，实验二数据中的粘滞打字现象更为少见 ($t_{15} = -6.32, p < .001$)。粘滞打字现象指的是在前一只手指抬起之前，下一只手指已经按下。在实验二中，被试的打字动作更慢，因此粘滞打字现象会减少^[196]。

如图5.7所示，实验者计数了各种误触类型。计数的方法是基于触点数量的，例如，五指休息导致的五个误触会被统计五次。所有误触类型中，最常见的三种误触分别是多指休息导致的误触 (82.90%)、小鱼际误触 (7.53%) 和轻触误触 (3.21%)。对于图中半透明展示的误触 (h) 单指休息误触和 (i) 重触误触，人类（实验者）无法在不知情上下文的情况下区分其有意性，因此实验者认为这两类误触对于机器来说也是难以区分的，它们加起来的占比是 1.52%。实验者用实验二的数据重新训练了误触分类器。留一（被试）交叉验证显示其准确率达到 98.88% ($SD=0.73\%$)。

量	介绍	实验一	实验二	Levene	T 检验
触摸压力	有意触摸的平均压力	188.39g SD=64.72	124.75g SD=60.71	拒绝	$t_{15} = -2.78$, $p < .01$
连续触摸	连续触摸数量占有意触摸的比例	4.02% SD=1.96%	11.89% SD=4.41%	通过	$t_{15} = 3.57$, $p < .005$
粘滞打字	粘滞打字数量占有意触摸的比例	17.60% SD=9.34%	7.73% SD=5.22%	通过	$t_{15} = -6.32$, $p < .001$

表 5.3 实验一和实验二之间的用户打字行为差异。本节用 T 检验来评估差异的显著性，如果 Levene 测试拒绝了数据的同方差性，本节将使用不等方差 T 检验来代替 T 检验。

假正率为 2.27% (SD=2.00%)，假负率为 0.65% (SD=0.63%)。平均而言，用户在每 100 次有意触摸中，只会遇到 0.93 次误触和 0.65 次未识别触摸，这说明该技术的防误触能力很强，而且能够轻易处理用户将手指休息在触摸屏上的情形。至此，本章完成了对终版防误触技术 (TypeBoard) 的设计。



图中展示了所有的九种误触的情形，其中括号中的数值是该类型误触占所有误触的比例。

图 5.7 所有的误触类型图示

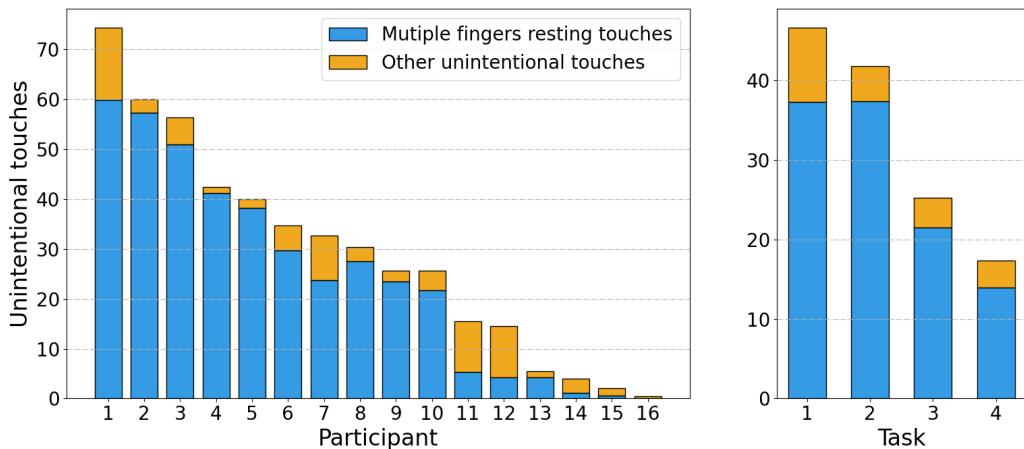
5.4.3 关于防误触键盘的讨论

5.4.3.1 与先前工作的对比

TapBoard^[163]是与本章最类似的相关工作，请注意本章所介绍的 TypeBoard 与 TapBoard 是不同的，这两份工作都探讨了触摸屏键盘上的误触问题。然而，先前工作 TapBoard 存在两个致命缺陷：首先，TapBoard 强行将敲击手势定为有意触摸，将其它触摸手势定为误触，其中敲击手势指的是位移小于 15 毫米、触摸时长短于 450 毫秒的触摸。这种要求人去适应技术的设计违反了人机交互的自然性原则。第二，TapBoard 只能在触摸的手指抬起时判断触摸的有意性，响应性比正常的键盘低，尽管如此，其识别准确率也只有 97%。当实验者复现 TapBoard 技术，并将其应用于本节实验所采集的自然打字行为数据集上时，TapBoard 的误触识别准确率仅为 73.83% ($SD=14.39\%$)，这一表现远低于本章所介绍的防误触技术 TypeBoard (98.88%， $SD = 0.73\%$)。以上对比表明，本章所介绍的防误触技术显著优于先前技术。

5.4.3.2 用户差异

如图5.8（左侧）所示，不同被试的打字行为有较大差异。举一些极端的例子，被试 P1 平均每 100 次有意触摸就会产生 74.40 次无意触摸，而被试 P16 在整个实验过程中仅产生了三次无意触摸。实验者按被试的心理将他们分为三类：（1）P1-P10 属于第一类被试，他们愿意在打字的过程中将非交互手指休息在触摸屏上，因此会产生大量的无意触摸。（2）P11-P12 属于第二类被试，他们相信 TypeBoard 能够防止误触，放心地将手掌休息在触摸屏的边缘，但很少将非交互手指休息在触摸屏上。系统为这部分被试过滤了大量的边缘误触。P12 评论道：“我不会将手指休息在触摸屏上，因为已经在触摸屏上的手指无法像在物理键盘上一样直接按下去。”（3）P13-P16 在打字过程中时刻将双手悬空，他们在 TypeBoard 上打字的行为和在普通平板电脑上打字的行为并无太大差别，说明他们并未主动尝试 TypeBoard 防误触键盘带来的新体验。P13-P16 在平板电脑键盘上的打字经验差异很大，P13 和 P14 在平板电脑上打字经验仅为一年，P15 没有相关经验。P16 是平板电脑打字的专家用户，他几乎不产生任何无意触摸，这是因为他再普通的平板电脑键盘上也能盲打。上述分析说明，用户个性化对于本章介绍的防误触键盘技术来说是很有价值的。



左图是不同被试在无意触摸计数上的差异，右图是不同任务在无意触摸计数上的差异。

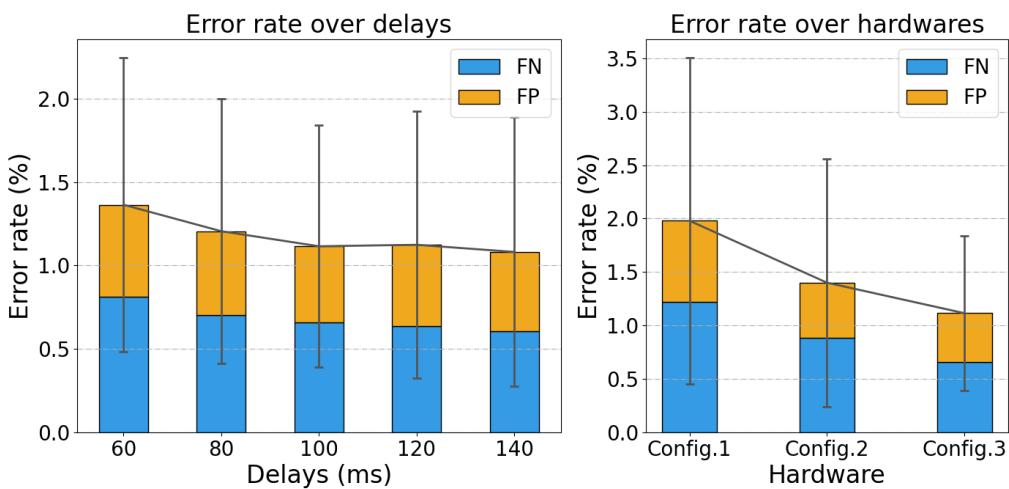
图 5.8 用户差异和任务差异图示

5.4.3.3 任务差异

如图5.8（右侧）所示，不同文本输入任务下无意触摸的数量有较大差异。在四个不同的实验任务下，被试平均每 100 次触摸会产生 46.67 次 ($SD=48.01$)、41.78 次 ($SD=52.65$)、25.28 次 ($SD=37.82$) 和 17.37 次 ($SD=14.50$) 无意触摸。匹配样本 T 检验显示以下实验任务之间，多指休息导致的无意触摸数量存在显著性差异：1-3 ($t_{15} = 2.58, p < .05$)、1-4 ($t_{15} = 2.23, p < .05$)。被试在第四个实验任务中更少地将手指休息在触摸板上，这是因为第四个任务看图写话不需要频繁切换文本输入和光标控制。结果表明，文本输入任务的多样性对误触研究来说是十分重要的，然而，这一多样性在相关文献中往往被忽视^[163,172,186]。

5.4.3.4 为什么要采样 100 毫秒的数据来判别误触？

数据采用的时间越长，有用的信息越多，对误触的判别准确率也会越高。然而，太长的采样时间又会导致较大的延迟，影响用户体验。因此，技术的识别准确率和延迟之间存在权衡的问题。为了找到一个好的平衡点，实验者通过模拟实验评测了不同延迟下误触识别准确率的变化，结果如图5.9（左侧）所示。重复测量方差分析显示延迟对误触识别准确率有显著影响 ($F_{4,56} = 4.88, p < 0.05$)，后验测试显示以下延迟对之间存在显著差异：60 毫秒-100 毫秒 ($p < .05$)、60 毫秒-120 毫秒 ($p < .05$)、60 毫秒-140 毫秒 ($p < .05$)、80 毫秒-140 毫秒 ($p < .05$)。也就是说，当延迟超过 100 毫秒时，延迟的进一步增长不再带来准确率的提高。因此，100 毫秒的识别延迟是最佳选择，使得误触识别的准确率达到 98.88%。触摸交互中 100 毫秒的延迟是用户可察觉的，但不会显著影响用户的交互效率^[197-199]。



左图展示了不同信号采集时间下误触识别的准确率，右图展示了不同信号组合下误触识别的准确率。

图 5.9 采样时间和信道对误触识别准确率的影响

5.4.3.5 本技术是否适用于更少的传感器？

本章的各种实验数据已经证明，压敏触摸板上的压力信号对误触识别的准确率来说是至关重要的。然而，目前市面上大多数触摸屏设备都不具备全屏幕压力传感的能力，而小部分设备（如 MacBook 触摸板）能够传感触摸板受到的总压力。为了探索本技术是否适用于更少的传感器，实验者通过数据模拟分析了三种硬件设置下本防误触技术的表现。

- 电容触摸屏：**最常用的电容式触摸屏设备仅能提供电容屏信号，而不能提供任何压力相关的信号。为了评测本防误触技术在电容式触摸屏上的表现，实验者删除了支持向量机模型所用特征向量中所有和压力信号有关系的特征，并重新训练模型。
- 压力触摸板：**苹果电脑的压力触摸板可以用四个角上的压力传感器计算触摸板所受到的总压力。为了评测本防误触技术在压力触摸板上的表现，实验者将特征向量中涉及到的触点压力信号用估算值来代替，触点压力的估算方法是用总压力乘以该触点的接触面积占所有触点接触面积之和的比例。
- 压敏触摸板：**本实验中用到的设备即为压敏触摸板，能够直接测量每个触点的压力。

如图5.9（右侧）所示是本防误触技术在上述三种硬件设置下的误触识别率。重复测量方差分析显示硬件设置对误触识别的准确率有显著性影响 ($F_{2,28} = 10.52, p < 0.001$)。Bonferroni 校正后验测试显示以下硬件设置对之间存在显著性差异：1-2 ($p < .005$)、1-3 ($p < .005$)。设置 2 和设置 3 之间的准确率差异仅为趋

势 ($p=.062$)。这一结果表明，本防误触技术在压敏触摸板上的表现最佳；而在只能识别总压力的压力触摸板上表现也不错，压力触摸板是性价比很高的选择。

5.4.3.6 迭代方法

由于技术会影响用户行为，而数据驱动的技术在研发过程中又需要先采集用户行为数据，所以数据驱动的技术和用户行为之间存在死锁问题。本章工作提出了解决上述死锁问题的迭代方法，即实验者首先研发了初版防误触技术，采集用户在初版防误触键盘上的打字行为数据，最后根据行为数据研发终版防误触技术。相比之下，大多数先前研究都在无反馈的情况下（类似本章的实验一）收集用户的行为数据^[174-175,177]，或者在替代性反馈下收集数据^[163,172]，以回避死锁问题。作者认为本章所述的迭代方法值得更多的关注，在本章所介绍的工作中，实验二采集到的数据和实验一数据之间存在较大差异，也就是说，迭代的方法能够得到更接近真实的用户行为数据，而基于更真实数据的交互技术也会更能理解人的交互意图。

5.4.3.7 为什么不采用深度学习？

本章使用经典的机器学习方法（支持向量机）来解决防误触问题，而没有用到深度学习，这是因为，防误触技术作为触摸屏设备的一项底层技术，需要速度快、功耗低的解决方案。深度学习是一种计算密集型的算法，并不满足防误触技术的要求。出于这个原因，作者认为经典的机器学习方法更适用于本技术。

5.4.3.8 提高防误触能力其它方法

首先，可以利用触点与键盘布局之间的关系辅助对触点有意性的判断，例如，当触点没有落在任何按钮上时，它更有可能是一次误触。第二，可以使用语音模型作为先验知识。贝叶斯解码器被广泛应用于从词汇表中预测用户所需的单词^[95,109,112,147]。贝叶斯解码器还可以用于计算后续触摸点落在不同位置上的概率分布，当触摸落在低概率的按键上时，它更有可能是误触。在本章介绍的防误触技术当中，实验者没有将键盘布局和语言模型作为先验知识，因为本章所探索的是面向平板电脑键盘的通用的防误触方法。

5.5 防误触键盘的评测

5.5.1 实验三：防误触键盘的评测实验

实验三的目的有两点：首先，实验三比较了平板电脑上防误触键盘 TypeBoard 和普通软键盘的文本输入效率；第二，如本章引言所述，若在防误触键盘上加上按键的轮廓反馈，将能大大提升文本输入效率，因此实验三还评测了 TypeBoard+ 的表现，TypeBoard+ 即为加上了按键轮廓反馈的防误触键盘。总的来说，本实验比较了被试在三种平板电脑键盘上的打字效率和用户体验：（1）普通触摸屏键盘；（2）TypeBoard；（3）TypeBoard+。

5.5.1.1 被试

实验者从校园中招募了 15 名被试参与实验，被试的年龄从 19 岁到 26 岁不等，平均年龄为 20.87 岁，标准差为 2.42，其中有 7 名女性被试。所有被试都是右撇子，他们没有参与过之前的用户实验。被试们拥有不少于两年的智能手机打字经验，平均经验为 6.67 年，标准差为 2.06。11 名被试有平板电脑十指打字经验。

5.5.1.2 实验设计和过程

本实验采用了组内实验设计，即每名被试都会使用三种不同的键盘进行实验。被试坐在一张可以调节高度的座椅上，他们可以将座椅调整至舒适的高度。被试们在一块压敏触摸板上打字，在平板电脑上获取视觉反馈。如图5.10所示，实验涉及到三种键盘设置：

1. **设置一：普通键盘。**在普通的平板电脑文本输入中，触摸板上所有的触摸都会被视为有意交互，被试在打字过程中需要双手悬空，或者将手腕休息在触摸板外，以避免误触。
2. **设置二：TypeBoard。**实验二中所介绍的平板电脑防误触键盘，键盘只会对它认为的有意触摸做出响应，被试在打字过程中可以将非交互手指休息在触摸板上。
3. **设置三：TypeBoard+。**在防误触键盘的基础上，加上按键轮廓的触觉反馈。为了提供触觉反馈，实验者在每个按键上贴上了 0.05 毫米厚的贴纸，在 F 键和 J 键上有额外的凸起，这一凸起的设置和物理键盘一样。被试可以通过触摸贴纸的边缘来定位自己的手指，从而有机会实现盲打。

在每种键盘设置下，有五段重复实验，每段实验被试誊写一段中文文字。誊写是文本输入研究中最常用的实验任务^[151,194-195]，一般用于评测文本输入法的打字效率上限。每段中文文本大约包含 100 个字，这些文本段落是从一个打字测速网



Ordinary keyboard

TypeBoard

TypeBoard plus

从左到右三幅子图分别表示实验三

图 5.10 实验三的实验设置

站^[200]中随机抽选的。被试需要尽可能又快又准地完成文本输入实验。实验采用拉丁方来平衡三种键盘设置的出场顺序。在每一种键盘设置的正式实验开始前，被试有五分钟的热身时间通过誊写文本来熟悉键盘的设置，热身过程中用过的文本不会出现在正式实验中。在每两段实验之间，被试休息两分钟的时间以避免疲劳。平均而言，本实验的总时长为 90 分钟。

5.5.2 防误触键盘的评测结果

实验者使用重复测量方差分析来评估组内因素键盘设置对打字速度、未纠正错误率（UER）和已纠正错误率（CER）的显著性影响。由于 UER 和 CER 不服从正态分布，实验者在方差分析之前使用对齐秩变换算法^[153]校准数据。如果有任何独立变量对实验结果有显著性影响 ($p < .05$)，实验者采用 Bonferroni 校正后的后验测试来做成对比较。

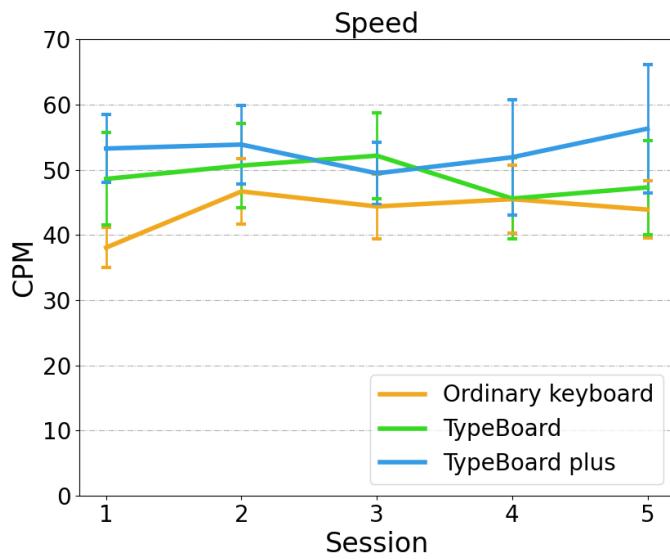
5.5.2.1 打字速度

本节将用每分钟输入中文字个数（CPM）来测量被试的打字速度：

$$CPM = \frac{|S|}{T} \times 60 \quad (5.1)$$

其中， $|S|$ 是所誊写的中文文本的长度（包括标点符号）， T 是被试完成输入的时间，即从第一次有意触摸到最后一次有意触摸之间所用的时间。

如图5.11所示是被试的打字速度随着实验段数变化的折线图。实验段数对打字速度没有显著性影响 ($F_{4,56} = 1.76, p = .15$)，结果表明，这三款平板电脑键盘的学习成本很低，用户刚上手就能达到可观的文本输入速度。键盘设置对打字速度有显著影响 ($F_{2,28} = 26.76, p < .001$)，成对比较显示，所有的键盘设置之间都存在显著差异：普通键盘-TypeBoard ($p < .005$)、普通键盘-TypeBoard+ ($p < .001$)、TypeBoard-TypeBoard+ ($p < .005$)。被试在普通键盘上的平均打字速度为 43.71



图中展示了三种不同键盘设置下，被试的打字速度随着实验时间推移而变化的折线图，误差条表示 95% 置信区间。

图 5.11 三种键盘下被试打字速度对比

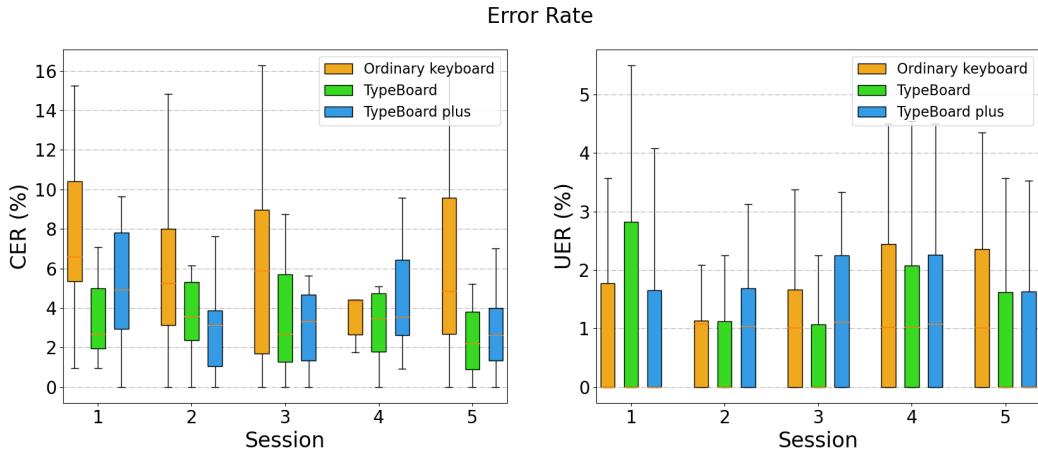
在 TypeBoard 上的打字速度为 48.87 CPM ($SD=10.14$)，超出普通键盘打字效率 11.78%。被试在 TypeBoard+ 上的打字速度为 52.97 CPM ($SD=9.85$)，超出普通键盘打字效率 21.19%。结果表明，面向平板电脑文本输入的防误触技术大幅度提高了打字效率。

为了比较 TypeBoard+ 和物理键盘，实验者组织了一次非正式实验测量了实验三被试在物理键盘上的打字速度。被试们通过自己熟悉的物理键盘在一个中文打字测速网站上输入五段文字^[200]，平均速度达到 65.01 CPM ($SD=9.26$)。这一结果表明，TypeBoard+ 和物理键盘之间还是有一定差距的，但是和普通平板电脑键盘相比，这一差距已经缩小了 43.48%。

5.5.2.2 打字错误率

本节用两种指标来评测文本输入法的错误率：(1) 未纠正错误率 (UER) —— 遗留在所誊写文本中的错误，UER 等于为未经纠错的中文字个数除以所誊写句子的字数；(2) 已纠正错误率 (CER) —— 那些在打字过程中被纠正 (例如通过删除) 的错误，CER 等于被纠正的中文字个数除以誊写句子的字数。被试使用拼音输入法过程中对字母的纠正未算入 CER 指标中。由于 UER 和 CER 不服从正态分布，实验者使用对齐秩变换算法^[153]校正数据。

如图5.12所示是已纠正错误率 (CER) 和未纠正错误率 (UER) 随实验段数变化的箱型图。实验段数对 CER 没有显著影响 ($F_{4,56} = 1.01, p = .39$)，键盘设置对



图中展示了三种键盘设置下，已纠正错误率和未纠正错误率的箱型图。

图 5.12 打字错误率图示

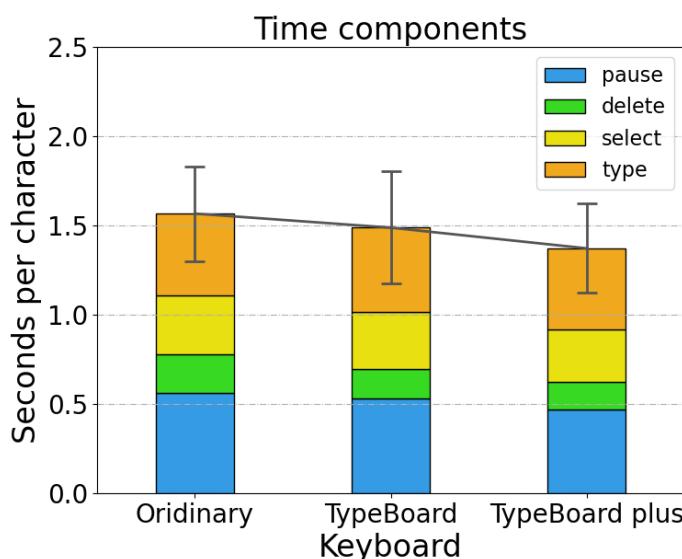
CER 有显著影响 ($F_{2,28} = 9.49, p < .005$)，后验测试显示以下键盘设置的 CER 之间存在显著差异：普通键盘-TypeBoard ($p < .01$)、普通键盘-TypeBoard+ ($p < .005$)。普通键盘、TypeBoard 和 TypeBoard+ 下的平均 CER 分别是 6.66% ($SD=4.42\%$)、4.58% ($SD=3.58\%$) 和 4.21% ($SD=2.58\%$)。实验段数 ($F_{4,56} = 0.41, p = .71$) 和键盘设置 ($F_{2,28} = 0.001, p = .998$) 都对 UER 没有显著影响。普通键盘、TypeBoard 和 TypeBoard+ 下的平均 UER 分别是 1.29% ($SD=1.67\%$)、1.28% ($SD=1.38\%$) 和 1.28% ($SD=1.16\%$)。结果表明，防误触技术降低了用户在平板电脑触摸屏上打字时打错字的概率，这是 TypeBoard 提高了平板电脑打字速度的主要原因。与 TypeBoard 相比，TypeBoard+ 在降低打字错误率上没有显著的优势，这说明 TypeBoard+ 的打字效率更高是另有原因的。

5.5.2.3 时间构成

为了深入探讨本技术的用户打字效率问题，实验者将被试的文本输入耗时拆分成三个构成部分：键入时间、选词时间和停顿时间。键入时间指的是被试点击单词各个字母所用的时间，是从被试点击单词首字母到点击单词尾字母之间的时间。选词时间是被试从候选词列表中选中所需单词的时间。停顿时间是从完成一个单词的选词到输入下一个单词首字母所需的时间。上述时间指的都是每输入一个英文单词所消耗的时间。

为了深入探讨三种键盘设置对用户打字效率和体验的影响，实验者将被试的文本输入耗时拆分成四个构成部分：键入时间、选词时间、删除时间和停顿时间。键入时间是被试点选拼音中每个字母所用的时间，是从被试点击拼音首字母到点击拼音尾字母之间的时间。选词时间是被试从拼音输入法中选择所需中文字或词

所用的时间。删除时间是删除拼音或中文字所需的时间。停顿时间是从完成一个中文字或词到输入下一个拼音首字母所需的时间。如图5.13所示是每输入一个中文字所消耗的键入时间、选词时间、删除时间和停顿时间。重复测量方差分析显示键盘设置对选词时间 ($F_{2,28} = 7.85, p < .005$)、删除时间 ($F_{2,28} = 20.89, p < .001$) 和停顿时间 ($F_{2,28} = 12.76, p < .001$) 都有显著影响。后验测试显示 TypeBoard ($p < .001$) 和 TypeBoard+ ($p < .001$) 都与普通键盘相比都能显著降低删除时间，这一结果表明，防误触键盘技术降低了打字的错误率。TypeBoard+ 在停顿时间上显著优于普通键盘 ($p < .001$) 和 TypeBoard ($p < .005$)，这一结果表明 TypeBoard+ 上的按键轮廓反馈让用户有机会实现盲打，省去了键入新字时目光在任务文本和键盘之间来回切换的时间。



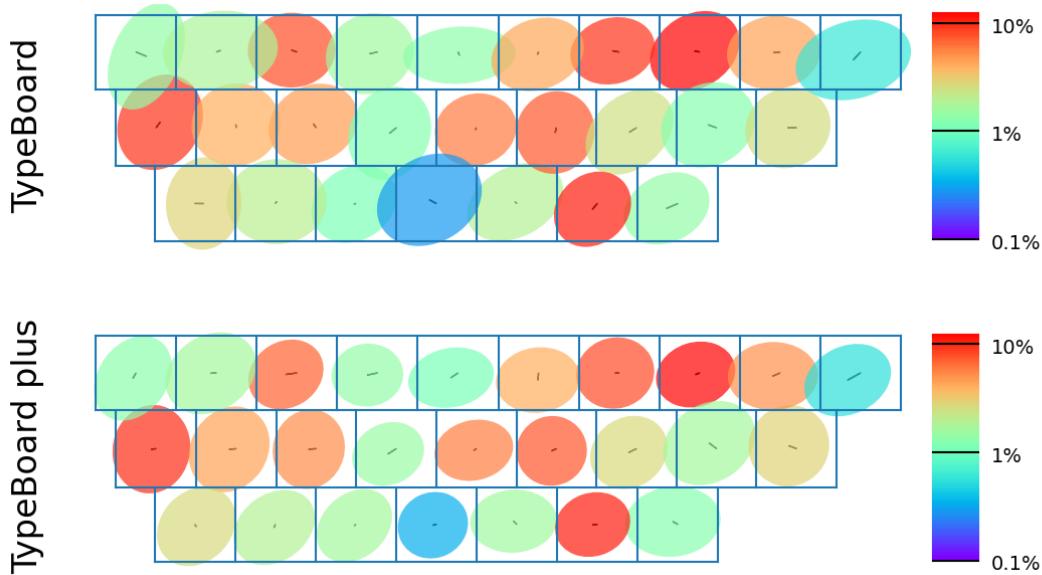
图中展示了被试每输入一个中文字时各个动作所消耗的时间，误差条表示 95% 置信区间

图 5.13 时间构成

5.5.2.4 触点位置

如图5.14所示分别是 TypeBoard 和 TypeBoard+ 下有意触摸位置的点云分布。为了方便观察，实验者默认点云符合二维高斯分布。实验中键盘每个按键的宽度为 17 毫米，TypeBoard 上触摸点云中心点位置与键盘中心位置之间在 X 轴/Y 轴上的偏移量分别为 -1.03/-0.29 毫米，TypeBoard+ 上偏移量为 -2.02/-0.59 毫米。这一偏移量相对于按键宽度来说是很小的，说明被试的触摸点击非常接近每个按键的中心。TypeBoard 上触摸点云在 X 轴/Y 轴上的标准差分别为 5.66/5.07 毫米，TypeBoard+ 上标准差为 5.01/4.53 毫米。成对采样 T 检验显示，键盘设置对触摸点云的在两个轴上标准差之积存在显著影响 ($t_{15} = -4.65, p < .001$)，这一结果表明，被试在

TypeBoard+ 上的触摸行为更加精准，TypeBoard+ 上的按键轮廓反馈有助于提高用户触摸行为的精准性，从而提高打字效率。



上图展示了 TypeBoard 下触摸点云的分布图，下图展示了 TypeBoard+ 的点云图。椭圆表示二维高斯分布的三倍标准差的范围。热度图表示每个字母被点击的频率。

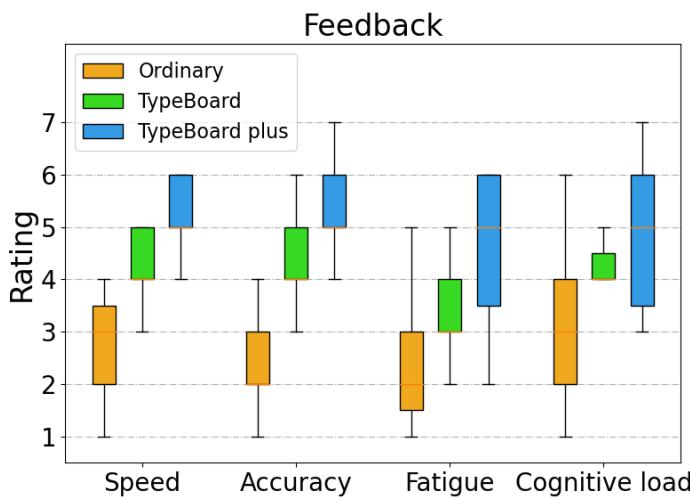
图 5.14 两种键盘的触摸点云图示

5.5.2.5 主观评分

被试在实验结束后，通过 7 级李克特量表为三种键盘的主观打字速度、打字准确率、疲劳程度和认知负担打分（1 分最差，7 分最佳）。如图 5.15 所示是主观评分的结果。Wilcoxon Signed-Rank 测试表明 TypeBoard 显著提升了平板电脑键盘的主观打字速度 ($Z = -2.27, p < .05$)、打字准确率 ($Z = -3.24, p < .005$)，降低了疲劳感 ($Z = -2.84, p < .005$) 和认知负担 ($Z = -1.99, p < .05$)。TypeBoard+ 也显著提升了平板电脑键盘的主观打字速度 ($Z = -3.17, p < .005$)、打字准确率 ($Z = -3.52, p < .001$)，降低了疲劳感 ($Z = -3.34, p < .001$) 和认知负担 ($Z = -2.28, p < .05$)。TypeBoard+ 在主观打字速度 ($Z = -2.40, p < .05$) 和疲劳感 ($Z = -2.85, p < .005$) 上比 TypeBoard 更优。结果表明 TypeBoard 和 TypeBoard+ 都优化了平板电脑触摸屏打字的用户体验。

5.5.2.6 实验结果小结

(1) 与平板电脑触摸屏上的普通键盘相比，TypeBoard 防误触技术将打字效率提高了 11.78%。具有缓解疲劳、降低认知负担和减少打字错误率的好处。(2) TypeBoard+ 进一步将 TypeBoard 的打字效率提升 8.51%，与普通键盘相比提高了



图中展示了三种键盘设置下被试的主观打字速度、打字准确率、疲劳程度和认知负担打分（1分最差，7分最佳）。

图 5.15 主观评分

21.19%。TypeBoard+ 的优点是提高用户打字行为的精准性和减少打字过程中寻找按键的时间。结果表明，本触摸屏防误触技术提高了文本输入技术的效率和用户体验，也为触摸屏上的盲打体验奠定了基础。

5.6 本章小结

本章介绍了面向平板电脑文本输入的防误触技术（TypeBoard），它可以过滤98.88%的误触。该技术强大的防误触能力改变了用户在平板电脑上打字的行为习惯，使得用户在平板电脑上打字时愿意将非交互手指休息在触摸屏上，以避免长时间打字带来的疲劳。平均而言，用户在 TypeBoard 上每 100 次有意触摸就会产生 40.83 次无意触摸，大部分都被该防误触技术过滤了。评测实验发现，防误触技术将平板电脑上的文本输入效率提高了 11.78%。在用户可以将手指休息在触摸屏上，而不引发误触的前提下，实验者在触摸屏上加上了键盘按键的轮廓反馈，使得用户可以在触摸屏上盲打，这进一步将平板电脑上的文本输入效率提高到原来的 21.18%。结果显示，防误触技术在诸多方面提升了平板电脑文本输入法的效率和用户体验，包括缓解疲劳、降低打字错误率和间接支持了触摸屏盲打。

第6章 结论和展望

触摸交互是人机交互的重要组成部分，是人控制手指接触交互表面，通过点击、长按、滑动等手势向计算机输入信息的方式。

6.1 结论

6.2 展望

参考文献

- [1] Jota R, Ng A, Dietz P, et al. How fast is fast enough? a study of the effects of latency in direct-touch pointing tasks[C]//Proceedings of the sigchi conference on human factors in computing systems. 2013: 2291-2300.
- [2] Lam A H, Li W J, Liu Y, et al. Mids: micro input devices system using mems sensors[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems: volume 2. IEEE, 2002: 1184-1189.
- [3] Alsop T. Shipment forecast of tablets, laptops and desktop pcs worldwide from 2010 to 2024 (in million units)[J]. Statista, 2020.
- [4] Hololens2[EB/OL]. 2021. <https://www.microsoft.com/en-us/hololens>.
- [5] Gu Y, Yu C, Li Z, et al. Accurate and low-latency sensing of touch contact on any surface with finger-worn imu sensor[C]//Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology. 2019: 1059-1070.
- [6] Jakobsen M R, Jansen Y, Boring S, et al. Should i stay or should i go? selecting between touch and mid-air gestures for large-display interaction[C]//IFIP Conference on Human-Computer Interaction. Springer, 2015: 455-473.
- [7] Miller R B. Response time in man-computer conversational transactions[C]//Proceedings of the December 9-11, 1968, fall joint computer conference, part I. 1968: 267-277.
- [8] Ng A, Lepinski J, Wigdor D, et al. Designing for low-latency direct-touch input[C]//Proceedings of the 25th annual ACM symposium on User interface software and technology. 2012: 453-464.
- [9] Yi X, Yu C, Shi W, et al. Is it too small?: Investigating the performances and preferences of users when typing on tiny qwerty keyboards[J]. International Journal of Human-Computer Studies, 2017, 106: 44-62.
- [10] Harrison C, Benko H, Wilson A D. Omnitouch: wearable multitouch interaction everywhere [C]//Proceedings of the 24th annual ACM symposium on User interface software and technology. 2011: 441-450.
- [11] Xiao R, Schwarz J, Throm N, et al. Mrtouch: Adding touch input to head-mounted mixed reality [J]. IEEE transactions on visualization and computer graphics, 2018, 24(4): 1653-1660.
- [12] Paradiso J A, Hsiao K y, Strickon J, et al. Sensor systems for interactive surfaces[J]. IBM Systems Journal, 2000, 39(3.4): 892-914.
- [13] Agarwal A, Izadi S, Chandraker M, et al. High precision multi-touch sensing on surfaces using overhead cameras[C]//Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'07). IEEE, 2007: 197-200.
- [14] Chang J S, Kim E Y, Jung K, et al. Real time hand tracking based on active contour model [C]//International Conference on Computational Science and Its Applications. Springer, 2005: 999-1006.

-
- [15] Letessier J, Bérard F. Visual tracking of bare fingers for interactive surfaces[C]//Proceedings of the 17th annual ACM symposium on User interface software and technology. 2004: 119-122.
 - [16] Sugita N, Iwai D, Sato K. Touch sensing by image analysis of fingernail[C]//2008 SICE Annual Conference. IEEE, 2008: 1520-1525.
 - [17] Grudin J. Integrating paper and digital information on enhanceddesk: a method for realtime finger tracking on an augmented desk system[J]. ACM Transactions on Computer-Human Interaction (TOCHI), 2001, 8(4): 307-322.
 - [18] Saba E N, Larson E C, Patel S N. Dante vision: In-air and touch gesture sensing for natural surface interaction with combined depth and thermal cameras[C]//2012 IEEE International Conference on Emerging Signal Processing Applications. IEEE, 2012: 167-170.
 - [19] Xiao R, Hudson S, Harrison C. Direct: Making touch tracking on ordinary surfaces practical with hybrid depth-infrared sensing[C]//Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces. 2016: 85-94.
 - [20] Benko H, Jota R, Wilson A. Miragetable: freehand interaction on a projected augmented reality tabletop[C]//Proceedings of the SIGCHI conference on human factors in computing systems. 2012: 199-208.
 - [21] Wilson A D, Benko H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces[C]//Proceedings of the 23rd annual ACM symposium on User interface software and technology. 2010: 273-282.
 - [22] Newcombe R A, Izadi S, Hilliges O, et al. Kinectfusion: Real-time dense surface mapping and tracking[C]//2011 10th IEEE international symposium on mixed and augmented reality. IEEE, 2011: 127-136.
 - [23] Mistry P, Maes P. Mouseless: a computer mouse as small as invisible[M]//CHI'11 Extended Abstracts on Human Factors in Computing Systems. 2011: 1099-1104.
 - [24] Xiao R, Harrison C, Hudson S E. Worldkit: rapid and easy creation of ad-hoc interactive applications on everyday surfaces[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013: 879-888.
 - [25] Chan L, Liang R H, Tsai M C, et al. Fingerpad: private and subtle interaction using fingertips [C]//Proceedings of the 26th annual ACM symposium on User interface software and technology. 2013: 255-260.
 - [26] Chen K Y, Lyons K, White S, et al. utrack: 3d input using two magnetic sensors[C]//Proceedings of the 26th annual ACM symposium on User interface software and technology. 2013: 237-244.
 - [27] Parizi F S, Whitmire E, Patel S. Auraring: Precise electromagnetic finger tracking[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2019, 3(4): 1-28.
 - [28] Shi Y, Zhang H, Zhao K, et al. Ready, steady, touch! sensing physical contact with a finger-mounted imu[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4(2): 1-25.
 - [29] Gu Y, Yu C, Li Z, et al. Qwertyring: Text entry on physical surfaces using a ring[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4(4): 1-29.

-
- [30] Meier M, Strel P, Fender A, et al. Tapld: Rapid touch interaction in virtual reality using wearable sensing[C]//2021 IEEE Virtual Reality and 3D User Interfaces (VR). IEEE, 2021: 519-528.
 - [31] Oh J Y, Lee J, Lee J H, et al. Anywheretouch: Finger tracking method on arbitrary surface using nailed-mounted imu for mobile hmd[C]//International Conference on Human-Computer Interaction. Springer, 2017: 185-191.
 - [32] Niikura T, Watanabe Y, Ishikawa M. Anywhere surface touch: utilizing any surface as an input area[C]//Proceedings of the 5th Augmented Human International Conference. 2014: 1-8.
 - [33] Liu G, Gu Y, Yin Y, et al. Keep the phone in your pocket: Enabling smartphone operation with an imu ring for visually impaired people[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4(2): 1-23.
 - [34] Downs R. Using resistive touch screens for human/machine interface[J]. Analog Applications Journal Q, 2005, 3: 5-10.
 - [35] Hecht D S, Thomas D, Hu L, et al. Carbon-nanotube film on plastic as transparent electrode for resistive touch screens[J]. Journal of the Society for information Display, 2009, 17(11): 941-946.
 - [36] Lee S, Buxton W, Smith K. A multi-touch three dimensional touch-sensitive tablet[J]. Acm Sigchi Bulletin, 1985, 16(4): 21-25.
 - [37] Wang F, Ren X. Empirical evaluation for finger input properties in multi-touch interaction [C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2009: 1063-1072.
 - [38] Han J Y. Low-cost multi-touch sensing through frustrated total internal reflection[C]//Proceedings of the 18th annual ACM symposium on User interface software and technology. 2005: 115-118.
 - [39] Matsushita N, Rekimoto J. Holowall: designing a finger, hand, body, and object sensitive wall [C]//Proceedings of the 10th annual ACM symposium on User interface software and technology. 1997: 209-210.
 - [40] Wilson A D. Touchlight: an imaging touch screen and display for gesture-based interaction [C]//Proceedings of the 6th international conference on Multimodal interfaces. 2004: 69-76.
 - [41] Paradiso J A, Leo C K, Checka N, et al. Passive acoustic sensing for tracking knocks atop large interactive displays[C]//SENSORS, 2002 IEEE: volume 1. IEEE, 2002: 521-527.
 - [42] Xiao R, Lew G, Marsanico J, et al. Toffee: enabling ad hoc, around-device interaction with acoustic time-of-arrival correlation[C]//Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services. 2014: 67-76.
 - [43] Heo S, Lee G. Forcetap: extending the input vocabulary of mobile touch screens by adding tap gestures[C]//Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services. 2011: 113-122.
 - [44] [M].
 - [45] Kane S K, Avrahami D, Wobbrock J O, et al. Bonfire: a nomadic system for hybrid laptop-tabletop interaction[C]//Proceedings of the 22nd annual ACM symposium on User interface software and technology. 2009: 129-138.

-
- [46] Masson D, Gogey A, Malacia S, et al. Whichfingers: identifying fingers on touch surfaces and keyboards using vibration sensors[C]//Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology. 2017: 41-48.
 - [47] Stearns L, Oh U, Findlater L, et al. Touchcam: Realtime recognition of location-specific on-body gestures to support users with visual impairments[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018, 1(4): 1-23.
 - [48] Fitts P M. The information capacity of the human motor system in controlling the amplitude of movement.[J]. Journal of experimental psychology, 1954, 47(6): 381.
 - [49] MacKenzie I S. Movement time prediction in human-computer interfaces[M]//Readings in human-computer interaction. Elsevier, 1995: 483-493.
 - [50] Wobbrock J O, Cutrell E, Harada S, et al. An error model for pointing based on fitts' law [C]//Proceedings of the SIGCHI conference on human factors in computing systems. 2008: 1613-1622.
 - [51] Ko Y J, Zhao H, Kim Y, et al. Modeling two dimensional touch pointing[C]//Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology. 2020: 858-868.
 - [52] Bi X, Li Y, Zhai S. Fitts law: modeling finger touch with fitts' law[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013: 1363-1372.
 - [53] Bützler J, Vetter S, Jochems N, et al. Bivariate pointing movements on large touch screens: Investigating the validity of a refined fitts' law[J]. Work, 2012, 41(Supplement 1): 3526-3532.
 - [54] El Lahib M, Tekli J, Issa Y B. Evaluating fitts' law on vibrating touch-screen to improve visual data accessibility for blind users[J]. International Journal of Human-Computer Studies, 2018, 112: 16-27.
 - [55] Vetter S, Bützler J, Jochems N, et al. Fitts' law in bivariate pointing on large touch screens: Age-differentiated analysis of motion angle effects on movement times and error rates[C]// International Conference on Universal Access in Human-Computer Interaction. Springer, 2011: 620-628.
 - [56] Xia H, Jota R, McCanny B, et al. Zero-latency tapping: using hover information to predict touch locations and eliminate touchdown latency[C]//Proceedings of the 27th annual ACM symposium on User interface software and technology. 2014: 205-214.
 - [57] Uno Y, Kawato M, Suzuki R. Formation and control of optimal trajectory in human multijoint arm movement[J]. Biological cybernetics, 1989, 61(2): 89-101.
 - [58] Flash T, Hogan N. The coordination of arm movements: an experimentally confirmed mathematical model[J]. Journal of neuroscience, 1985, 5(7): 1688-1703.
 - [59] Galloway J C, Koshland G F. General coordination of shoulder, elbow and wrist dynamics during multijoint arm movements[J]. Experimental brain research, 2002, 142(2): 163-180.
 - [60] Leigh D, Forlines C, Jota R, et al. High rate, low-latency multi-touch sensing with simultaneous orthogonal multiplexing[C]//Proceedings of the 27th annual ACM symposium on User interface software and technology. 2014: 355-364.
 - [61] Gu Y, Yu C, Chen X, et al. Typeboard: Identifying unintentional touch on pressure-sensitive touchscreen keyboards[C]//The 34th Annual ACM Symposium on User Interface Software and Technology. 2021: 568-581.

- [62] Xu X, Yu C, Wang Y, et al. Recognizing unintentional touch on interactive tabletop[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4 (1): 1-24.
- [63] Le H V, Kosch T, Bader P, et al. Palmtouch: Using the palm as an additional input modality on commodity smartphones[C]//Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018: 1-13.
- [64] Nelson W L. Physical principles for economies of skilled movements[J]. Biological cybernetics, 1983, 46(2): 135-147.
- [65] Welford A T. Fundamentals of skill.[J]. 1968.
- [66] Kopper R, Bowman D A, Silva M G, et al. A human motor behavior model for distal pointing tasks[J]. International journal of human-computer studies, 2010, 68(10): 603-615.
- [67] Kopp R E. Pontryagin maximum principle[M]//Mathematics in Science and Engineering: volume 5. Elsevier, 1962: 255-279.
- [68] Lewis F L, Vrabie D, Syrmos V L. Optimal control[M]. John Wiley & Sons, 2012.
- [69] Ross I M, Proulx R J, Karpenko M. An optimal control theory for the traveling salesman problem and its variants[J]. arXiv preprint arXiv:2005.03186, 2020.
- [70] Kamien M I, Schwartz N L. Dynamic optimization: the calculus of variations and optimal control in economics and management[M]. courier corporation, 2012.
- [71] Ross I M, Karpenko M, Proulx R J. A nonsmooth calculus for solving some graph-theoretic control problems[J]. IFAC-PapersOnLine, 2016, 49(18): 462-467.
- [72] Zhang Z. Microsoft kinect sensor and its effect[J]. IEEE multimedia, 2012, 19(2): 4-10.
- [73] Humpherys J, Redd P, West J. A fresh look at the kalman filter[J]. SIAM review, 2012, 54(4): 801-823.
- [74] Lewis F L, Lewis F. Optimal estimation: with an introduction to stochastic control theory[M]. Wiley New York, 1986.
- [75] Bertsekas D P. Nonlinear programming[J]. Journal of the Operational Research Society, 1997, 48(3): 334-334.
- [76] Žilinskas A. Practical mathematical optimization: An introduction to basic optimization theory and classical and new gradient-based algorithms[M]. JSTOR, 2006.
- [77] Bertsimas D, Tsitsiklis J. Simulated annealing[J]. Statistical science, 1993, 8(1): 10-15.
- [78] Conn A R, Gould N I, Toint P L. Trust region methods[M]. SIAM, 2000.
- [79] The sensel morph[J/OL]. Morph Technology, 2021. <https://morph.sensel.com/>.
- [80] Keselman L, Iselin Woodfill J, Grunnet-Jepsen A, et al. Intel realsense stereoscopic depth cameras[C]//Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017: 1-10.
- [81] Point N. Optitrack[J]. Natural Point, Inc, 2011.
- [82] Madgwick S. An efficient orientation filter for inertial and inertial/magnetic sensor arrays[J]. Report x-io and University of Bristol (UK), 2010, 25: 113-118.

-
- [83] Cao X, Wilson A D, Balakrishnan R, et al. Shapetouch: Leveraging contact shape on interactive surfaces[C]//2008 3rd IEEE international workshop on horizontal interactive human computer systems. IEEE, 2008: 129-136.
 - [84] Harrison C, Schwarz J, Hudson S E. Tapsense: enhancing finger interaction on touch surfaces [C]//Proceedings of the 24th annual ACM symposium on User interface software and technology. 2011: 627-636.
 - [85] Ramos G, Boulos M, Balakrishnan R. Pressure widgets[C]//Proceedings of the SIGCHI conference on Human factors in computing systems. 2004: 487-494.
 - [86] Heo S, Lee G. Force gestures: augmented touch screen gestures using normal and tangential force[M]//CHI'11 Extended Abstracts on Human Factors in Computing Systems. 2011: 1909-1914.
 - [87] Roudaut A, Lecolinet E, Guiard Y. Microrolls: expanding touch-screen input vocabulary by distinguishing rolls vs. slides of the thumb[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2009: 927-936.
 - [88] Xiao R, Schwarz J, Harrison C. Estimating 3d finger angle on commodity touchscreens[C]// Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces. 2015: 47-50.
 - [89] Badger P. Capacitive sensing library-arduino playground[J]. 2018.
 - [90] De Smedt Q, Wannous H, Vandeborre J P. Skeleton-based dynamic hand gesture recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2016: 1-9.
 - [91] Ge L, Cai Y, Weng J, et al. Hand pointnet: 3d hand pose estimation using point sets[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 8417-8426.
 - [92] Spurr A, Song J, Park S, et al. Cross-modal deep variational hand pose estimation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 89-98.
 - [93] Yuan S, Garcia-Hernando G, Stenger B, et al. Depth-based 3d hand pose estimation: From current achievements to future goals[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 2636-2645.
 - [94] Yu C, Gu Y, Yang Z, et al. Tap, dwell or gesture? exploring head-based text entry techniques for hmds[C]//Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017: 4479-4488.
 - [95] Gupta A, Ji C, Yeo H S, et al. Rotoswype: Word-gesture typing using a ring[C]//Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019: 1-12.
 - [96] Grossman T, Chen X A, Fitzmaurice G. Typing on glasses: Adapting text entry to smart eyewear [C]//Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services. 2015: 144-152.
 - [97] Yu C, Sun K, Zhong M, et al. One-dimensional handwriting: Inputting letters and words on smart glasses[C]//Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016: 71-82.

- [98] Lu Y, Yu C, Yi X, et al. Blindtype: Eyes-free text entry on handheld touchpad by leveraging thumb's muscle memory[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2017, 1(2): 1-24.
- [99] Kienzle W, Hinckley K. Lightring: always-available 2d input on any surface[C]//Proceedings of the 27th annual ACM symposium on User interface software and technology. 2014: 157-160.
- [100] Markussen A, Jakobsen M R, Hornbæk K. Selection-based mid-air text entry on large displays [C]//IFIP Conference on Human-Computer Interaction. Springer, 2013: 401-418.
- [101] Schick A, Morlock D, Amma C, et al. Vision-based handwriting recognition for unrestricted text input in mid-air[C]//Proceedings of the 14th ACM international conference on Multimodal interaction. 2012: 217-220.
- [102] Oney S, Harrison C, Ogan A, et al. Zoomboard: a diminutive qwertype soft keyboard using iterative zooming for ultra-small devices[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013: 2799-2802.
- [103] Yi X, Yu C, Xu W, et al. Compass: Rotational keyboard on non-touch smartwatches[C]// Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017: 705-715.
- [104] Chen X, Grossman T, Fitzmaurice G. Swipeboard: a text entry technique for ultra-small interfaces that supports novice to expert transitions[C]//Proceedings of the 27th annual ACM symposium on User interface software and technology. 2014: 615-620.
- [105] Hong J, Heo S, Isokoski P, et al. Splitboard: A simple split soft keyboard for wristwatch-sized touch screens[C]//Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015: 1233-1236.
- [106] Gupta A, Balakrishnan R. Dualkey: miniature screen text entry via finger identification[C]// Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016: 59-70.
- [107] Leiva L A, Sahami A, Catala A, et al. Text entry on tiny qwertype soft keyboards[C]//Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015: 669-678.
- [108] Shibata T, Afergan D, Kong D, et al. Driftboard: A panning-based text entry technique for ultra-small touchscreens[C]//Proceedings of the 29th Annual Symposium on User Interface Software and Technology. 2016: 575-582.
- [109] Gong J, Xu Z, Guo Q, et al. Wristtext: One-handed text entry on smartwatch using wrist gestures [C]//Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018: 1-14.
- [110] Ni T, Bowman D, North C. Airstroke: bringing unistroke text entry to freehand gesture interfaces[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011: 2473-2476.
- [111] Kristensson P O, Zhai S. Shark2: a large vocabulary shorthand writing system for pen-based computers[C]//Proceedings of the 17th annual ACM symposium on User interface software and technology. 2004: 43-52.

- [112] Markussen A, Jakobsen M R, Hornbæk K. Vulture: a mid-air word-gesture keyboard[C]// Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014: 1073-1082.
- [113] Amma C, Georgi M, Schultz T. Airwriting: Hands-free mobile text input by spotting and continuous recognition of 3d-space handwriting with inertial sensors[C]//2012 16th International Symposium on Wearable Computers. IEEE, 2012: 52-59.
- [114] Zhou S, Dong Z, Li W J, et al. Hand-written character recognition using mems motion sensing technology[C]//2008 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. IEEE, 2008: 1418-1423.
- [115] Devoe D B. Alternatives to handprinting in the manual entry of data[J]. IEEE Transactions on Human Factors in Electronics, 1967(1): 21-32.
- [116] Bajer B, MacKenzie I S, Baljko M. Huffman base-4 text entry glove (h4 teg)[C]//2012 16th International Symposium on Wearable Computers. IEEE, 2012: 41-47.
- [117] Goldstein M, Chincholle D. Finger-joint gesture wearable keypad[C]//second workshop on human computer interaction with mobile devices. 1999.
- [118] Peshock A, Duvall J, Dunne L E. Argot: A wearable one-handed keyboard glove[C]// Proceedings of the 2014 ACM international symposium on wearable computers: adjunct program. 2014: 87-92.
- [119] Rosenberg R, Slater M. The chording glove: a glove-based text input device[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 1999, 29(2): 186-191.
- [120] Seibel R. Performance on a five-finger chord keyboard.[J]. Journal of Applied Psychology, 1962, 46(3): 165.
- [121] Whitmire E, Jain M, Jain D, et al. Digitouch: Reconfigurable thumb-to-finger input and text entry on head-mounted displays[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2017, 1(3): 1-21.
- [122] Gummesson J, Priyantha B, Liu J. An energy harvesting wearable ring platform for gestureinput on surfaces[C]//Proceedings of the 12th annual international conference on Mobile systems, applications, and services. 2014: 162-175.
- [123] Zhang C, Waghmare A, Kundra P, et al. Fingersound: Recognizing unistroke thumb gestures using a ring[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2017, 1(3): 1-19.
- [124] Zhang C, Xue Q, Waghmare A, et al. Fingerping: Recognizing fine-grained hand poses using active acoustic on-body sensing[C]//Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018: 1-10.
- [125] Kim J, Delamare W, Irani P. Thumbtext: Text entry for wearable devices using a miniature ring [C]//Graphics Interface. 2018.
- [126] Xu Z, Wong P C, Gong J, et al. Tiptext: eyes-free text entry on a fingertip keyboard[C]// Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology. 2019: 883-899.

- [127] Nirjon S, Gummesson J, Gelb D, et al. Mobicys 2015 - proceedings of the 13th annual international conference on mobile systems, applications, and services[J]. 2015.
- [128] Partridge K, Chatterjee S, Sazawal V, et al. Tilttype: accelerometer-supported text entry for very small devices[C]//Proceedings of the 15th annual ACM symposium on User interface software and technology. 2002: 201-204.
- [129] Zhong M, Yu C, Wang Q, et al. Forceboard: Subtle text entry leveraging pressure[C]// Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018: 1-10.
- [130] Kurauchi A, Feng W, Joshi A, et al. Eyeswipe: Dwell-free text entry using gaze paths[C]// Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016: 1952-1956.
- [131] Majaranta P, Räihä K J. Text entry by gaze: Utilizing eye-tracking[J]. Text entry systems: Mobility, accessibility, universality, 2007: 175-187.
- [132] Sarcar S, Panwar P, Chakraborty T. Eyek: an efficient dwell-free eye gaze-based text entry system[C]//Proceedings of the 11th asia pacific conference on computer human interaction. 2013: 215-220.
- [133] Shilkrot R, Huber J, Steimle J, et al. Digital digits: A comprehensive survey of finger augmentation devices[J]. ACM Computing Surveys (CSUR), 2015, 48(2): 1-29.
- [134] Gupta A, Yang J, Balakrishnan R. Asterisk and obelisk: Motion codes for passive tagging[C]// Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology. 2018: 725-736.
- [135] Ogata M, Sugiura Y, Osawa H, et al. Iring: intelligent ring using infrared reflection[C]// Proceedings of the 25th annual ACM symposium on User interface software and technology. 2012: 131-136.
- [136] Tsukada K, Yasumura M. Ubi-finger: A simple gesture input device for mobile and ubiquitous environment[J]. Journal of Asian Information, Science and Life (AISL), 2004, 2(2): 111-120.
- [137] Van Kleek M G, Styke W, Schraefel M, et al. Finders/keepers: a longitudinal study of people managing information scraps in a micro-note tool[C]//Proceedings of the SIGCHI conference on human factors in computing systems. 2011: 2907-2916.
- [138] Zhang C, Wang X, Waghmare A, et al. Fingorbites: interaction with wearables using synchronized thumb movements[C]//Proceedings of the 2017 ACM International Symposium on Wearable Computers. 2017: 62-65.
- [139] Pratichizzo D, Chinello F, Pacchierotti C, et al. Remotouch: A system for remote touch experience[C]//19th International Symposium in Robot and Human Interactive Communication. IEEE, 2010: 676-679.
- [140] Solazzi M, Frisoli A, Bergamasco M. Design of a novel finger haptic interface for contact and orientation display[C]//2010 IEEE Haptics Symposium. IEEE, 2010: 129-132.
- [141] Ashbrook D, Baudisch P, White S. Nenya: subtle and eyes-free mobile input with a magnetically-tracked finger ring[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011: 2043-2046.

- [142] Harrison C, Hudson S E. Abracadabra: wireless, high-precision, and unpowered finger input for very small mobile devices[C]//Proceedings of the 22nd annual ACM symposium on User interface software and technology. 2009: 121-124.
- [143] Nguyen A, Banic A. 3dtouch: A wearable 3d input device for 3d applications[M]. IEEE, 2015.
- [144] Zhang C, Xue Q, Waghmare A, et al. Soundtrak: Continuous 3d tracking of a finger using active acoustics[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2017, 1(2): 1-25.
- [145] Chan L, Chen Y L, Hsieh C H, et al. Cyclopsring: Enabling whole-hand and context-aware interactions through a fisheye ring[C]//Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology. 2015: 549-556.
- [146] Yang X D, Grossman T, Wigdor D, et al. Magic finger: always-available input through finger instrumentation[C]//Proceedings of the 25th annual ACM symposium on User interface software and technology. 2012: 147-156.
- [147] Gordon M, Ouyang T, Zhai S. Watchwriter: Tap and gesture typing on a smartwatch miniature keyboard with statistical decoding[C]//Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016: 3817-3821.
- [148] Goodman J, Venolia G, Steury K, et al. Language modeling for soft keyboards[C]//Proceedings of the 7th international conference on Intelligent user interfaces. 2002: 194-195.
- [149] Ide N. The american national corpus: Then, now, and tomorrow[C]//Selected Proceedings of the 2008 HCSNet Workshop on Designing the Australian National Corpus: Mustering Languages, Summerville, MA. Cascadilla Proceedings Project. 2008.
- [150] Azenkot S, Zhai S. Touch behavior with different postures on soft smartphone keyboards[C]// Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services. 2012: 251-260.
- [151] MacKenzie I S, Soukoreff R W. Phrase sets for evaluating text entry techniques[C]//CHI'03 extended abstracts on Human factors in computing systems. 2003: 754-755.
- [152] Findlater L, Wobbrock J O, Wigdor D. Typing on flat glass: examining ten-finger expert typing patterns on touch surfaces[C]//Proceedings of the SIGCHI conference on Human factors in computing systems. 2011: 2453-2462.
- [153] 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.
- [154] Arif A S, Stuerzlinger W. Analysis of text entry performance metrics[C]//2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH). IEEE, 2009: 100-105.
- [155] Grandjean E, Kroemer K H. Fitting the task to the human: a textbook of occupational ergonomics[M]. CRC press, 1997.
- [156] Nirjon S, Gummesson J, Gelb D, et al. Typingring: A wearable ring platform for text input[C]// Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services. 2015: 227-239.

- [157] Royal S, Zhai S, Kristensson P O. Performance and user experience of touchscreen and gesture keyboards in a lab setting and in the wild[C]//Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM, 2015: 679-688.
- [158] Yajima T, Hosobe H. A japanese software keyboard for tablets that reduces user fatigue[C]// 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC): volume 1. IEEE, 2018: 339-346.
- [159] Kim J H, Aulck L, Bartha M C, et al. Differences in typing forces, muscle activity, comfort, and typing performance among virtual, notebook, and desktop keyboards[J]. Applied ergonomics, 2014, 45(6): 1406-1413.
- [160] Bonner M N, Brudvik J T, Abowd G D, et al. No-look notes: Accessible eyes-free multi-touch text entry[C]//International Conference on Pervasive Computing. Springer, 2010: 409-426.
- [161] Tinwala H, MacKenzie I S. Eyes-free text entry with error correction on touchscreen mobile devices[C]//Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries. 2010: 511-520.
- [162] Sears A. Improving touchscreen keyboards: design issues and a comparison with other devices [J]. Interacting with computers, 1991, 3(3): 253-269.
- [163] Kim S, Son J, Lee G, et al. Tapboard: making a touch screen keyboard more touchable[C]// Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013: 553-562.
- [164] Crump M J, Logan G D. Warning: This keyboard will deconstruct—the role of the keyboard in skilled typewriting[J]. Psychonomic bulletin & review, 2010, 17(3): 394-399.
- [165] Gordon A M, Soechting J F. Use of tactile afferent information in sequential finger movements [J]. Experimental brain research, 1995, 107(2): 281-292.
- [166] Logan G D, Crump M J. Hierarchical control of cognitive processes: The case for skilled typewriting[M]//Psychology of learning and motivation: volume 54. Elsevier, 2011: 1-27.
- [167] Ma Z, Edge D, Findlater L, et al. Haptic keyclick feedback improves typing speed and reduces typing errors on a flat keyboard[C]//2015 IEEE World Haptics Conference (WHC). IEEE, 2015: 220-227.
- [168] Tactus technology[J/OL]. Tactus Technology, 2020. <http://www.tactustechnology.com/>.
- [169] Amberg M, Giraud F, Semail B, et al. Stimitac: a tactile input device with programmable friction [C]//Proceedings of the 24th annual ACM symposium adjunct on User interface software and technology. 2011: 7-8.
- [170] Bau O, Poupyrev I, Israr A, et al. Teslatouch: electrovibration for touch surfaces[C]// Proceedings of the 23rd annual ACM symposium on User interface software and technology. 2010: 283-292.
- [171] Levesque V, Oram L, MacLean K, et al. Enhancing physicality in touch interaction with programmable friction[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011: 2481-2490.
- [172] Xu X, Yu C, Wang Y, et al. Recognizing unintentional touch on interactive tabletop[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4 (1): 1-24.

- [173] Lu H, Li Y. Gesture on: Enabling always-on touch gestures for fast mobile access from the device standby mode[C]//Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015: 3355-3364.
- [174] Matero J, Colley A. Identifying unintentional touches on handheld touch screen devices[C]// Proceedings of the designing interactive systems conference. 2012: 506-509.
- [175] Annett M, Anderson F, Bischof W F, et al. The pen is mightier: understanding stylus behaviour while inking on tablets[M]//Proceedings of Graphics Interface 2014. 2014: 193-200.
- [176] Wang J, Zhao K, Zhang X, et al. Ubiquitous keyboard for small mobile devices: harnessing multipath fading for fine-grained keystroke localization[C]//Proceedings of the 12th annual international conference on Mobile systems, applications, and services. 2014: 14-27.
- [177] Le H V, Kosch T, Bader P, et al. Palmtouch: Using the palm as an additional input modality on commodity smartphones[C]//Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018: 1-13.
- [178] Durojaiye O M, Abzarian D. Classification of touch input as being unintended or intended[M]. Google Patents, 2016.
- [179] Kaikuranta T. Method for preventing unintended touch pad input due to accidental touching [M]. Google Patents, 2006.
- [180] Rigazio L, Kryze D, Morin P, et al. System and method for differentiating between intended and unintended user input on a touchpad[M]. Google Patents, 2013.
- [181] Brakensiek J, Bose R. Method and apparatus for precluding operations associated with accidental touch inputs[M]. Google Patents, 2013.
- [182] Popescu V. Touch screen with unintended input prevention[M]. Google Patents, 2015.
- [183] Kim S, Lee G. Tapboard 2: Simple and effective touchpad-like interaction on a multi-touch surface keyboard[C]//Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016: 5163-5168.
- [184] Griffin J T. Touch screen palm input rejection[M]. Google Patents, 2013.
- [185] Annett M, Gupta A, Bischof W F. Exploring and understanding unintended touch during direct pen interaction[J]. ACM Transactions on Computer-Human Interaction (TOCHI), 2014, 21(5): 1-39.
- [186] Schwarz J, Xiao R, Mankoff J, et al. Probabilistic palm rejection using spatiotemporal touch features and iterative classification[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014: 2009-2012.
- [187] Le H V, Mayer S, Steuerlein B, et al. Investigating unintended inputs for one-handed touch interaction beyond the touchscreen[C]//Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services. 2019: 1-14.
- [188] Goyins G S, Resman M F. Palm pressure rejection method and apparatus for touchscreens[M]. Google Patents, 2001.
- [189] Chaparro B, Nguyen B, Phan M, et al. Keyboard performance: ipad versus netbook[J]. Usability News, 2010, 12(2): 1-9.

- [190] Zhang Y, Pahud M, Holz C, et al. Sensing posture-aware pen+ touch interaction on tablets[C]// Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019: 1-14.
- [191] Matulic F, Vogel D, Dachselt R. Hand contact shape recognition for posture-based tabletop widgets and interaction[C]//Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces. 2017: 3-11.
- [192] Koura S, Suo S, Kimura A, et al. Amazing forearm as an innovative interaction device and data storage on tabletop display[C]//Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces. 2012: 383-386.
- [193] Fallot-Burghardt W, Fjeld M, Speirs C, et al. Touch&type: a novel pointing device for notebook computers[C]//Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles. 2006: 465-468.
- [194] Soukoreff R W, MacKenzie I S. Metrics for text entry research: An evaluation of msd and kspc, and a new unified error metric[C]//Proceedings of the SIGCHI conference on Human factors in computing systems. 2003: 113-120.
- [195] Yi X, Yu C, Shi W, et al. Word clarity as a metric in sampling keyboard test sets[C]//Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017: 4216-4228.
- [196] Dhakal V, Feit A M, Kristensson P O, et al. Observations on typing from 136 million keystrokes [C]//Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018: 1-12.
- [197] Attig C, Rauh N, Franke T, et al. System latency guidelines then and now—is zero latency really considered necessary?[C]//International Conference on Engineering Psychology and Cognitive Ergonomics. Springer, 2017: 3-14.
- [198] Kaaresoja T, Brewster S, Lantz V. Towards the temporally perfect virtual button: touch-feedback simultaneity and perceived quality in mobile touchscreen press interactions[J]. ACM Transactions on Applied Perception (TAP), 2014, 11(2): 1-25.
- [199] Kaaresoja T J. Latency guidelines for touchscreen virtual button feedback[D]. University of Glasgow, 2016.
- [200] Typing speed measurement.[J/OL]. Typing Speed Measurement., 2011. <http://www.51dzt.com/test/>.

附录 A 补充内容

附录是与论文内容密切相关、但编入正文又影响整篇论文编排的条理和逻辑性的资料，例如某些重要的数据表格、计算程序、统计表等，是论文主体的补充内容，可根据需要设置。

A.1 图表示例

A.1.1 图

附录中的图片示例（图 A.1）。

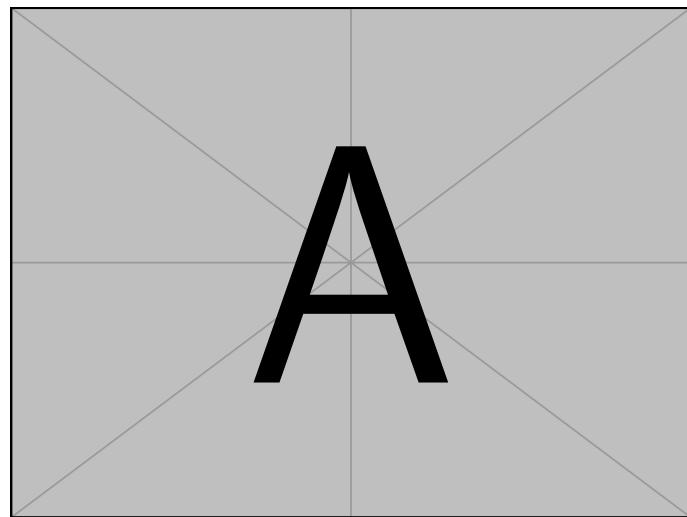


图 A.1 附录中的图片示例

致 谢

衷心感谢导师史元春教授和喻纯副教授对本人的精心指导。

感谢实验室同学

感谢家人朋友

感谢未婚妻

声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立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外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名： _____ 日 期： _____

个人简历、在学期间完成的相关学术成果

个人简历

1995年5月12日出生于广东省中山市

2013年9月考入清华大学计算机系计算机专业，2017年7月本科毕业并获得学士学位。

2017年9月保研进入清华大学计算机系攻读人机交互专业博士至今。

在学期间完成的相关学术成果

学术论文：

第一作者发表论文

- [1] **Gu Y**, Yu C, Chen X, et al. Typeboard: Identifying unintentional touch on pressure-sensitive touchscreen keyboards[C]//The 34th Annual ACM Symposium on User Interface Software and Technology. 2021: 568-581. (TH-CPL Rank A)
- [2] **Gu Y**, Yu C, Li Z, et al. Qwertyring: Text entry on physical surfaces using a ring[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4(4): 1-29. (TH-CPL Rank A)
- [3] **Gu Y**, Yu C, Li Z, et al. Accurate and low-latency sensing of touch contact on any surface with finger-worn imu sensor[C]//Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology. 2019: 1059-1070. (TH-CPL Rank A)

共同一作发表论文

- [4] Liu G, **Gu Y**, Yin Y, et al. Keep the phone in your pocket: Enabling smartphone operation with an imu ring for visually impaired people[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4(2): 1-23. (TH-CPL Rank A)

学生一作发表论文

- [5] Yu C, **Gu Y**, Yang Z, et al. Tap, dwell or gesture? exploring head-based text entry techniques for hmds[C]//Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017: 4479-4488. (TH-CPL Rank A)

非第一作者发表论文

- [6] Yu C, Wei X, Vachher S, Qin Y, Liang C, Weng Y, **Gu Y**, et al. Handsee:

enabling full hand interaction on smartphone with front camera-based stereo vision[C]//Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019: 1-13. (TH-CPL Rank A)

专利：

- [7] 史元春, 喻纯, 古裔正. 2021.07. 一种抬起手势的识别方法、系统、电子设备及存储介质. CN111580664A. (中国专利申请号, 已授权)
- [8] 史元春, 喻纯, 秦岳, 古裔正, 韦笑颖. 2021.07. 智能电子设备上基于单摄像头的双目视觉与物体识别技术. CN109993059B. (中国专利申请号, 已授权)
- [9] 喻纯, 古裔正, 杨志灿, 阎裕康, 史元春. 2018.08. 手型跟踪装置. CN207752443U. (实用新型, 已授权)
- [10] 史元春, 喻纯, 古裔正, 周诚驰, 张磊. 2022.01. 一种二维码扫描方法、装置及电子设备. CN113935348A. (中国专利公开号, 已公开)
- [11] 史元春, 喻纯, 古裔正, 周诚驰, 张磊. 2022.01. 一种控制智能家电的方法、装置及电子设备. CN113934150A. (中国专利公开号, 已公开)
- [12] 史元春, 喻纯, 古裔正. 2021.11. 触摸屏防误触的方法及装置、电子设备及存储介质. CN113608634A. (中国专利公开号, 已公开)
- [13] 史元春, 喻纯, 古裔正. 2020.08. 一种信息输入方法、系统、电子设备及存储介质. CN111580663A. (中国专利公开号, 已公开)
- [14] 史元春, 喻纯, 刘冠宏, 古裔正. 2020.08. 一种设备控制方法、电子设备、设备控制系统及存储介质. CN111580666A. (中国专利公开号, 已公开)

指导教师学术评语

论文提出了……

答辩委员会决议书

论文提出了……

论文取得的主要创新性成果包括：

1.
2.
3.

论文工作表明作者在xxxxx具有xxxxx知识，具有xxxx能力，论文xxxx，
答辩xxxx。

答辩委员会表决，（×票/一致）同意通过论文答辩，并建议授予×××（姓名）
×××（门类）学博士/硕士学位。