

神经设计学课程作业

——眼动软键盘设计

姓名： 张瑞升

学号： 220426

指导老师： 牛亚峰

日期： 2023.01.08

目 录

第一章 引言	1
1.1 课题背景	1
1.2 组织框架	2
第二章 文献综述.....	3
2.1 眼动交互	3
2.1.1 眼球运动特点	3
2.1.2 眼动特征提取	4
2.1.3 眼动交互现有问题	5
2.2 基于眼动交互的软键盘设计案例	5
2.2.1 软键盘的命令触发方式.....	7
2.2.2 软键盘的布局形式	9
2.2.3 软键盘的反馈形式	10
2.2.4 软键盘的预测能力	11
2.3 基于眼动交互的软键盘设计原则	12
第三章 眼动打字软键盘设计方案.....	13
3.1 软键盘设计方案一	13
3.1.1 界面介绍.....	13
3.1.2 使用流程.....	14
3.2 软键盘设计方案二	16
3.2.1 界面介绍.....	16
3.2.2 使用流程.....	17
3.3 软键盘设计方案三	19
3.3.1 界面介绍.....	19
3.3.2 使用流程.....	20
第四章 总结	22
参考文献.....	23
附录 I 眼动打字软键盘界面案例	28

第一章 引言

1.1 课题背景

随着传感计算、人工智能等信息技术的飞速发展与智能交互设备的不断革新，人机交互界面逐渐由基于鼠标、键盘、触控板等物理通道输入的图形用户界面向基于眼睛视线、手势动作、自然语言等感官通道输入的自然用户界面转变。在人与生俱来的自然感官通道中，以眼睛信息为特征的视觉通道最为主要^[1]。一方面，人的眼睛作为感受器持续地感知外界信息，从而完成人对计算机的信息认知；另一方面，人的眼睛作为效应器表达着自身的交互意图，辅助完成人对计算机的指令传达。伴随眼动追踪技术的不断完善，基于视觉通道的眼控交互^[2]逐渐成为自然人机交互系统中的重要交互方式之一，在辅助操作^[3]、人机协作^[4]、工效评价^[5]、自然驾驶^[6]、战场作战^[7]等场景得到了广泛应用。

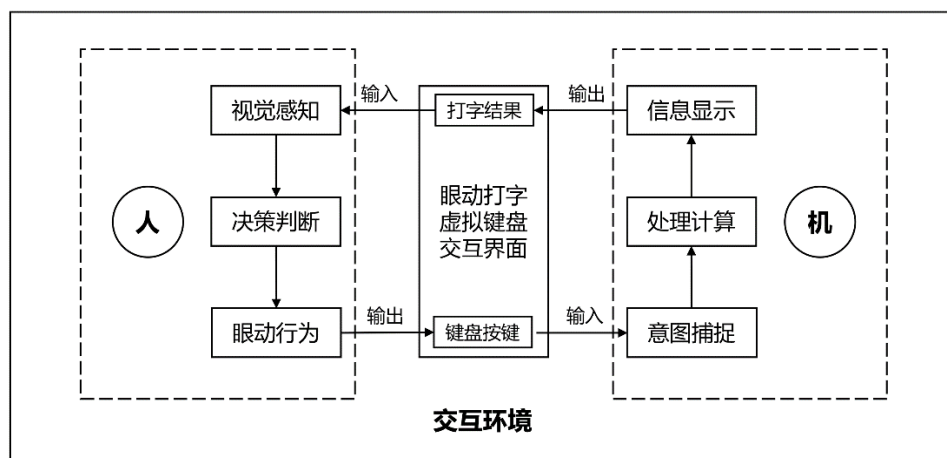


图 1-1 眼动打字行为模型

在眼控交互的目标人群中，以渐冻症人群为代表的严重运动障碍患者占据着庞大的比例^[8]。这类患者部分或全部丧失运动能力，甚至声带受损无法正常发声，只有眼睛能够灵活移动，因此视觉通道可能是其与外界交流的唯一途径。在该情况下，基于眼控交互的文字输入是其向他人表达自我意图的最主要方式。眼动打字行为模型如图 1-1 所示，用户通过有意识地注视软键盘按键并被计算机有效捕捉，计算机通过信息加工判断用户文字输入意图并将打字结果反馈给用户，从而

完成眼动打字过程。然而，由于眼睛在打字过程中同时作为感受器与效应器而不可避免地造成视觉通道过载问题^[9]，眼睛注视过程中存在漂移、震颤、无意识眼跳等视线抖动问题^[10]，以及眼动交互过程中的米达斯接触^[11]等问题，眼动打字仍然是一个缓慢低效的过程。基于人的视觉通道特点，设计合理的眼动打字软键盘布局形式与眼动打字交互方式，是建立舒适高效的眼动打字交互系统、提升特殊人群交互体验的关键环节。

1.2 组织框架

本报告共包含四章，各章节组织框架如图 1-2 所示。

第一章为引言，主要阐释眼动打字软键盘的课题背景与本文组织框架。

第二章为文献综述，主要综述了眼动交互特点，从命令触发、界面布局、反馈形式、预测能力四个方面整理了基于眼动交互的软键盘设计相关研究，并最终总结了基于眼动交互软键盘的普适设计原则。

第三章为设计方案阐释，共提出三种不同类型软键盘设计方案，并介绍每种方案的应用场景、界面布局、使用流程。

第四章为总结，针对本次课程报告进行整体总结。

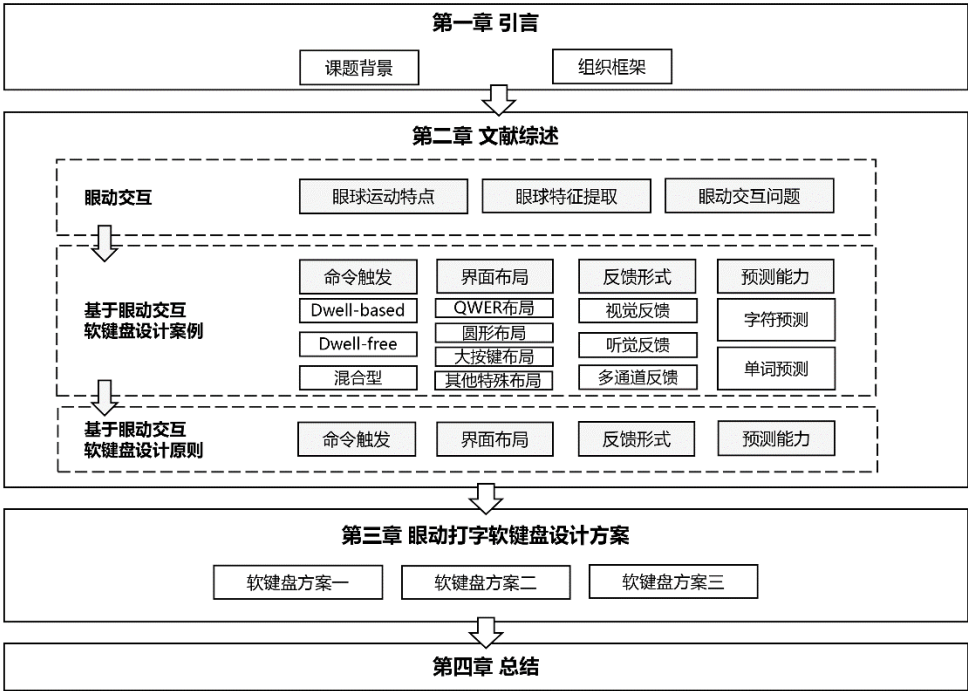


图 1-2 本报告组织框架

第二章 文献综述

人眼是一种精细敏感且动态变化的复杂生物学系统^[12]，其生理特征决定了眼动信息的模糊性。相应地，以人眼信息为输入模态的眼动打字交互系统必须充分贴合模糊输入下的眼球运动特点，设计符合人眼行为习惯的软键盘交互界面，以实现精确计算机交互系统与模糊眼动输入模态间的合理映射，进而提升眼动打字效率。本章为文献综述部分，主要阐释了眼动交互相关的人眼运动生理特点、眼动信息提取方法及眼动交互的现有问题，整理了不同形式的眼动打字软键盘交互界面设计案例，并总结了眼动打字软键盘交互界面设计原则。

2.1 眼动交互

眼动交互是指利用眼动追踪技术提取人的眼球运动特征数据及其对应的视觉注意行为，获取用户当前的视觉注意焦点、瞳孔扩张程度等时空参数，对用户视觉感知和认知活动进行分析推理，从而为交互任务提供数据输入和控制输出。本节将主要阐释眼动交互相关的眼球运动特点、特征提取方法及其现有问题。

2.1.1 眼球运动特点

眼睛是人视觉感知与意图表达的重要通道。正如眼心假说^[13]所证实，人的眼睛关注位置与其思考具有强关联性，用户直接的眼球运动表达着个体的交互意图。人的视觉形成源于视网膜上的感光细胞^[14]。在交互任务执行过程中，由于视网膜感光细胞的不均匀分布，个体必须要将交互目标落在视网膜中央凹区域上以最大限度提取交互目标的细节，因此人的眼睛必须通过不断移动来观察交互环境并搜索交互目标，通过注视停留来保持对交互目标的持续关注^[15]。快速性与动态性是眼球运动的重要生理特征。

眼球运动依照运动特点可以分为注视、扫视（也称眼跳）、眨眼、平滑追踪等基本运动状态。注视运动表示眼睛将中央凹视野区域保留在交互目标一定时间之上，来获取更多信息或触发交互操作，其被广泛应用于点选任务的选择触发环节^[16]。扫视运动表示人在观察交互目标过程中有意识或无意识的视线转移，其中：

无意识的视线转移被称为眼跳，其起点与终点未知，依赖于人的视觉注意；有意识的视线转移被称为眼势，依赖于人的交互意图；在普适交互任务中，眼势区分了用户是否存在交互意图，有效避免了米达斯接触问题，且相较于注视运动更为快速灵活，在近几年被频繁应用于指令多样的交互情境^[17]。眨眼运动是人自然状态下的正常生理行为，正常人眨眼时长一般为 200-400ms，部分研究人员将更长时间的有意识眨眼作为触发条件应用于眼动交互任务中^[18]。平滑追踪运动是一种连续反馈的状态，常出现于用户对交互环境内动态目标的持续关注，一般不用于主动情形下的眼动交互。基于上述基本眼动状态，设计对应的眼动交互策略与界面布局，以实现用户交互意图的识别与眼动交互指令触发，是建立眼动交互系统的普遍环节。

2.1.2 眼动特征提取

眼动交互中的眼动特征提取方法基于现代眼动追踪系统，一般采用计算机视觉技术完成眼球运动追踪与特征数据采集，其基本原理为：向人眼发射红外线，通过分析人眼对红外光线的反射状况确定人的视线方向，同时通过图像处理记录人眼瞳孔、虹膜、巩膜等视觉特征变化^[19]。Kar 等^[20]总结了眼动追踪系统所采用的具体算法。基于眼动追踪系统的使用环境，现代眼动追踪系统可分为桌面式眼动追踪系统与可穿戴式眼动追踪系统两类^[21]：桌面式眼动追踪系统将眼动仪固定放置在距离用户一定距离（如桌面）处观测眼球运动，常用于驾驶模拟等固定实验场景；可穿戴式眼动追踪系统将眼动仪集成在眼镜或头盔上，常用于虚拟现实等移动实验场景。根据眼球运动状态，眼动追踪系统可记录的具体眼动特征整理如表 2-1。

表 2-1 常用眼动交互特征

眼球运动状态	眼动交互特征具体参数
注视运动	注视持续时间、注视位置
扫视运动	扫视持续时间、扫视幅度、扫视潜伏期、扫视速度与方向
眨眼运动	眨眼持续时间，眨眼频率，眼睑闭合度
平滑追踪	追踪速度、追踪方向
其他	瞳孔大小、瞳孔扩张收缩程度

2.1.3 眼动交互现有问题

近年来,随着人工智能技术的飞速发展,以眼动追踪技术为基础的眼动交互准确性与交互效率不断提高。然而,受制于人眼运动特点,眼动交互仍然无法实现人与计算机实时、精确、舒适、高效的互动体验,存在着不可避免或难以解决的交互问题。在眼动系统交互设计中,解决、规避或合理利用此类问题,将显著提升眼动交互系统的可用性。眼动交互现有交互问题总结如下:

1.米达斯接触问题。米达斯接触^[11]又称为点石成金,指计算机错误识别用户因自然状态下随意性视线运动而产生的无意识意图,进而引发错误交互响应问题。米达斯接触本质是一种误操作,其解决根本在于区分用户明确的有意识意图与随意的无意识意图^[22]。目前,针对米达斯接触问题所广泛采用的解决思路有:采用较长时间的注视交互^[23];选择有意识的眨眼交互^[24];配合其他交互通道^[25]。

2.交互疲劳问题。人感知到的外界信息 80%以上从视觉通道获取^[1],在此基础上,人的眼睛若是继续执行过于复杂的交互任务,尤其要求用户长时间注视某个区域,则势必引起疲劳^[26]。针对交互疲劳问题,解决思路有两方面:在单通道交互方面,可采用有意识的扫视代替注视,或降低注视时间;在多通道交互方面,可由其他通道辅助视线交互。

3.空间准确度与交互延迟。主流眼动追踪设备识别精度在 0.5 度至 1 度之间,且空间识别准确度易受环境光、空间位置等因素影响^[27];由于眼动追踪设备基于摄像头识别眼动信息,因相机帧速率造成的识别延迟不可避免。交互精度的限制与延迟的存在要求眼动交互系统中交互元素应保持一定大小^[28],以应对客观层面精度延迟误差与主观层面视线抖动误差所带来的交互干扰。

2.2 基于眼动交互的软键盘设计案例

面向特殊人群的眼动打字交互系统要求实现更准确、更高效的打字体验,在现有交互技术基础上,应针对上一节所描述的眼动交互特点设计合理的软键盘交互界面。本节将从眼动交互软键盘的命令触发方式、布局形式、反馈形式与预测能力等方面综述现有眼动交互软键盘的界面设计案例。本节所提及的软键盘界面设计案例根据文献序号顺序整理如表 2-2 所示,各案例软键盘外观截图见附录 I。

表 2-2 本节所涉及的眼动交互软键盘设计案例

序号	界面/作者名	时间	命令触发	布局形式	反馈形式	预测能力	打字绩效
[32]	Frey	1990	Dwell-based	大按键	-	有	-
[33]	Lankford	2000	Dwell-based	QWER	缩小框	有	-
[34]	GazeTalk	2001	Dwell-based	大按键	缩小框	有	10WPM
[35]	pEYEs	2008	Dwell-based	圆形	颜色变化	有	7.85WPM
[36]	SliceType	2019	Dwell-based	圆形	颜色变化	有	-
[37]	Majaranta	2009	Dwell-based	QWER	进度条	无	17.3WPM
[38]	Mott	2017	Dwell-based	QWER	颜色变化	无	12.4WPM
[39]	Pi	2017	Dwell-based	QWER	进度条	无	18.4WPM
[40]	Pi	2020	Dwell-based	QWER	颜色变化	无	18.4WPM
[41]	Dasher	2000	Dwell-free	其他	字符长度	无	-
[43]	Quikwriting	2008	Dwell-free	圆形	-	无	5.0WPM
[44]	Eye-S	2009	Dwell-free	其他	-	无	6.8WPM
[45]	EyeSwipe	2016	Dwell-free	QWER	颜色变化	有	11.6WPM
[46]	Yu	2017	Dwell-free	QWER	颜色变化	有	5.03WPM
[47]	Liu	2016	Dwell-free	QWER	-	有	-
[48]	BlinkWrite	2011	Dwell-free	大按键	-	有	8WPM
[49]	Ringtext	2019	Dwell Free	圆形	视听反馈	有	8WPM
[50]	SMOOVS	2015	混合型	圆形	字符移动	无	3.34WPM
[51]	Zeng	2022	混合型	圆形	字符移动	无	4.7WPM
[53]	BigKey	2009	手机触摸	QWER	矩形框	有	23.6WPM
[54]	PressTapFlick	2022	Dwell Free	QWER	矩形框	无	11.4WPM
[56]	SPEye	2021	Dwell Free	圆形	字符移动	有	1WPM
[57]	Augkey	2016	Dwell-based	QWER	进度条	有	12WPM

2.2.1 软键盘的命令触发方式

眼动打字系统通过用户与屏幕前的虚拟软键盘进行眼动交互实现,在这一过程中,用户通过眼神指向需要交互的按键目标,通过注视或眼势等有意识的眼球动作完成软键盘的命令触发^[29]。因此,眼动打字是典型的指向选择过程,得益于眼球运动相较于其他指向方式具有更快速度的优势,眼动打字的效率主要取决于指向后的选择效率,即软键盘的命令触发^[30]。命令触发是眼动打字的基础,目前软键盘的命令触发方式主要分为两类:Dwell-based 与 Dwell-free。

1. Dwell-based 方式

Dwell-based 方式是一种基于注视的命令触发方式,也是眼动交互最早使用的命令触发方式。早在上世纪 80 年代,世界上第一台眼动交互计算机 Erica 便以注视运动作为计算机指令触发形式,Hutchinson 等^[31]详细描述了 Erica 的基本原理与其基于注视的眼动交互逻辑。Frey 等^[32]在 Erica 工作站技术上设计基于注视的眼动文字输入界面,受制于有限的注视点分辨率,该软键盘仅使用六个按键完成打字交互过程,并使用上下文敏感的预测模型完成字符预测,以提升打字速度。考虑到不同人所适应注视时间不尽相同,Lankford^[33]针对 Erica 打字系统提出了一种定制化的注视时间交互机制,并将传统 QWER 布局形式引入眼动打字软键盘设计中。21 世纪后,研究者将眼动打字推向更广泛场景。一款经典 Dwell-based 的预测性文本输入系统 GazeTalk^[34]被提出,其采用 750ms 注视时间,大按键布局满足了各类低分辨率眼动仪的交互需求。Huckauf 等^[35]将注视命令触发应用于饼状软件盘菜单,并采用 400ms 注视时间。Benligiray 等^[36]提出 SliceType 软键盘,其将初始注视时间设定为 1000ms,并允许用户根据熟练程度手动调低。

传统 Dwell-based 软键盘常采用 500-1000ms 注视时间作为触发条件,虽然缩短注视时间可以提高文本输入效率,但相应得会增加打字错误率。由于注视时间容易受到用户与上下文交互情境影响,部分研究者们开始探索自适应调节注视时间。Majaranta 等^[37]探究了用户对随时可调节注视时间的接收程度,其结果表明允许用户随意调节注视时间情况下用户打字绩效由 6.9 WPM 提升至 19.9 WPM。Mott 等^[38]提出级联注视打字方案,其注视时间随着用户文本输入过程而慢慢减小。Pi 等^[39-40]分别采用贝叶斯概率模型和动态贝叶斯网络根据过去用户输入的上下文动态调整不同按键触发所需的注视时间,以提升眼动打字速度。可调节的自

适应注视时间逐渐成为 Dwell-based 软键盘所广泛采用的重要优化方向。

2. Dwell-free 方式

注视时间的存在始终限制着输入速度，部分研究者开始采用其他形式眼球运动或其他通道代替注视完成眼动打字指令触发，这些免注视的眼动打字方式被称为 Dwell-free 方式。早期 Dwell-free 软键盘大都由连续手势指向交互中演变而来。例如：Ward 等^[41]提出了一种由连续二维手势驱动的文字输入界面——Dasher，其结合上下文来预测接下来输入字符的出现概率，根据概率高低显示字符区间大小，并根据当前手势指向进行补偿，直至字符最先超过基准线进而完成输入字符的选择；该方式使用连续指向来完成文字输入，大幅提升了文字输入效率，随后被集成 GazeTalk 软件中。Perlin^[42]提出了基于手写笔的触控文本输入方式 Quikwriting，该方式允许用户无需抬起便可连续书写；随后 Bee 等^[43]将 Quikwriting 迁移至眼动打字场景中，用户通过将视线由键盘内部区域迁移至外部区域完成字符组的选择，再由外部区域间迁移完成具体字符的选择，最终将视线由外部区域向内迁移完成字符键入。

除了由连续手势指向演变，部分免注视交互软键盘利用眼势或眨眼代替注视完成字符键入。例如：Porta 等^[44]提出类似手写板的 Eye-S 软键盘，其将交互屏幕分为九宫格，用户通过视线在各个格子间穿过绘制特定眼动轨迹形状完成字符键入；Kurauchi 等^[45]提出基于眼势的命令触发方法 EyeSwipe，该方法可以根据用户在 QWER 键盘上的注视路径推测输入单词，通过视线移动至键盘外再移动回来的眼势动作完成单词键入；Yu 等^[46]同样通过用户在 QWER 键盘上的扫描路径推测输入单词，但单词键入触发采用按键触发；Liu 等^[47]提出 LCSMapping 方法，通过用户在 QWER 键盘上的注视热图（AOI 分布）推理用户输入单词；Scott 等^[48]提出 BlinkWrite，采用有意识眨眼代替注视完成命令触发；Xu 等^[49]提出基于扇形菜单的 RingText 输入，每个扇形由两个字符组成，通过视线指向扇形区域后向上向下选择完成具体字符输入。然而，这些基于眼势的连续文本输入虽然极大程度提升输入速度，但存在错误识别的情况，整体准确率较低。

3. 混合型命令触发

部分研究者将更准确的 Dwell-based 和更快速的 Dwell-free 方式进行结合，以充分利用两种方式的优点。例如：Lutz 等^[50]提出了一种结合注视运动和平滑追

踪的混合输入软键盘 SMOOVS，将每四个字母上下左右排布成一簇，用户需要首先注视并追踪所要键入字母所在簇，随后字母簇随注视过程逐渐展开放大，最终用户通过平滑追踪最终目标字母完成字母选择；Zeng 等^[51]在 SMOOVS 的基础上对界面进行改进，字母簇在注视过程的中心位置不再变化。混合型软键盘输入方式有效避免交互过程中的 Midas 问题，并加快了文字输入速度，是眼动打字软键盘的可扩展方向。

2.2.2 软键盘的布局形式

现有眼动追踪技术将眼动交互精度限制在 0.5 度至 1 度视线范围，因此，眼动打字按键单元应限制在固定范围内。近年来，诸多研究者从调整眼动软键盘界面布局角度优化眼动打字，在有限的交互单元上扩展更为舒适的界面布局来改善用户体验。MacKenzie 等^[52]整理了不同类型的软键盘布局形式，并测试了六种布局形式的交互绩效：QWER、ABC 型、Dvorak 型、Fitaly 型、JustType 型和 Telephone 型。本文将软键盘的布局形式分为四类：为用户所熟悉的 QWER 布局、圆形布局、大按键布局与其他布局。

1. QWER 布局

QWER 布局作为最为用户熟悉的布局形式，可有效降低用户对界面布局的学习成本，是应用最为广泛的眼动打字界面布局之一。前文中所提及 Lankford^[33]、Majaranta 等^[37]、Mott 等^[38]、Pi 等^[39-40]软键盘均采用 QWER 的界面布局形式。然而，由于字符按键的密集分布和小字符单元的存在，QWER 界面眼动打字难度仍较高，故主要应用于基于眼动轨迹推理输入单词，例如 EyeSwipe^[45]。此外，部分研究者不满足于固定键盘形式，开始尝试动态变化的 QWER 界面布局。Al 等^[53]开发出自适应变化按键大小的移动设备软键盘 BigKey，其在传统 QWER 布局基础上根据输入字母状态预测接下来输入字符概率，根据概率大小分配按键布局；Rajanna 等^[54]将该原理应用于眼动界面布局中，并集成为基于视线指向脚踩选择的多模态打字方式 PressTapFlick。自适应调节的动态界面布局成为 QWER 界面布局的新趋势。

2. 圆形布局

由于人眼中央凹区域的周边视觉现象^[55]，围绕眼中心圆形分布的布局更贴合

人眼视觉特点,更有助于用户视觉搜索各按键的信息。许多研究者采用圆形布局形式,将字符键或字符簇分布在圆形区域上,以最大程度利用交互空间。Huckauf 等^[35]将 pEYES 饼状分层菜单应用与眼动交互场景,探究圆形的饼状菜单在打字任务和桌面导航任务上的交互效果; SliceType^[36]软键盘根据字符输入预测概率将字符以不同大小扇形的形状组成圆形键盘; QuikWriting^[43]将圆形键盘划分为 8 个区域,每个区域包含若干字符; Ringtext^[49]将外圆键盘完全分解,每个小扇形代表一个交互字符; Zeng 等^[51]参考 SMOOVS 将若干字符集成字符簇,六向圆形分布于视野边缘; Porta 等^[56]提出 SPEye,将字符簇十向分布,通过眼睛由中心向边缘平滑追踪完成字符簇选择。圆形布局有效提升了各方向视觉空间的利用率,是最贴合人视觉特点的布局方式。

3. 大按键布局

在眼动打字研究早期阶段,由于各类眼动仪分辨率低,因此眼动界面多采用大按键的布局形式,以最大限度减少交互错误。例如早期阶段 Frey^[32]受制于眼动追踪设备有限分辨率,仅采用六个大按键 3×2 布局完成字符组合与单词键入; GazeTalk^[34]通过布满屏幕的大按键布局构成沉浸式打字体验; Scott 等^[48]将 SAK 设计理念应用于眨眼输入中,采用四个大按键完成单词输入。然而,随着眼动追踪精度的不断提高,大按键布局容易产生更多字符组合过程的单词歧义,将字符分配在更多更小按键上成为主要趋势。

4. 其他布局

近年来,部分研究者跳出传统的 QWERTY 布局与眼中心布局,提出了一些新奇的眼动打字界面设计方案。其中,最为著名的特殊打字布局为 Dasher^[41],作为一种连续输入界面,其将上下文语言环境与用户输入意图结合,通过有意识调节字符区间长度完成连续字符键入,用户只需指向目标字符无需额外触发动作即可完成交互;部分研究人员将 Dasher 与 GazeTalk 进行组合,实现独特的眼动打字连续输入体验。

2.2.3 软键盘的反馈形式

打字反馈给予给用户交互引导与当前交互状态的直观感受,允许用户辨别当前交互是否符合自身预期,是提升眼动打字软键盘用户体验的关键环节。由于视

觉感知是人类信息感知的最主要通道，目前眼动打字软键盘反馈集中于视觉反馈。例如：Lankford^[33]提出一种缩小框的视觉反馈，通过缩小框缩小状态来提醒用户当前注视目标与注视持续时间，此类通过缩小框的打字意图可视化方式被广泛应用于 GazeTalk^[34]、PressTapFlick^[54]等 Dwell-based 软键盘中；Yu^[46]将色彩视觉反馈与视觉光标加入眼动打字软键盘中，以明确用户当前的指向目标，SliceType^[36]、RingText^[49]同样应用此类色彩反馈；Diaz-Tula^[57]提出 Augkey，在 QWER 按键布局基础上向目标按键附近添加单词预测显示，通过把视觉信息限制在中央凹区域，以最大限度减少眼球运动。

部分研究者在视觉反馈基础上添加其他通道的反馈，以缓解用户视觉通道过载问题。RingText^[49]在色彩反馈基础上加入了选择后的声音反馈；Zeng 等^[51]通过实验验证出单纯视觉反馈下的眼动打字交互绩效低于视听双通道反馈，然而，其实验结果表明部分用户仍然更偏好单一视觉反馈，表示声音的存在会分散注意力。多通道的反馈是目前减轻视觉通道压力的主要方向，然而各通道反馈需注意反馈力度，以免干扰用户交互状态。

2.2.4 软键盘的预测能力

根据用户已输入字符推理用户的期望单词是现代键盘中所广泛采用的环节，强大的软键盘预测能力有效简化了用户字符键入次数，加快了交互速度。例如，Benligiray 等^[36]在 SliceType 中将预测单词显示在用户当前键入按键附近，若符合用户期望单词，用户可以通过双重注视直接键入；Dasher^[41]将预测功能融合在字符可视化布局中，基于上文输入预测的字符概率将直接影响各字符的区间长度；MacKenzie^[58]提出基于单词、字母的眼动打字预测算法，除了基本单词的预测外，基于上文输入将接下来最可能输入的三个字符突出显示，以最大程度减小用户对交互目标的视觉搜索。软键盘界面中显示适度的预测信息有效简化交互流程，但值得注意的是，过量的预测信息显示将分散用户注意力。如何将预测信息有效布局在界面中是眼动打字软键盘设计需要考虑的重要问题。

2.3 基于眼动交互的软键盘设计原则

通过对现有眼动打字软键盘在命令触发、布局形式、反馈形式、预测能力等方面设计特点整理，结合眼动交互的基本特点，本节将基于眼动交互的软键盘设计原则总结如下：

一、命令触发方面：Dwell-based 方式具有稳定的命令触发能力，但注视时间的存在限制了交互速度，适用于需要稳定选择的交互情境，也可设计自适应注视时间的交互策略；Dwell-free 方式具有高灵活度与快速性的特点，适用于普适计算环境下的快速交互场景；混合型命令触发通过将 Dwell-based 和 Dwell-free 结合至打字动作的各个行为阶段，充分发挥两种交互方式特点，是值得深入探索的研究方向。

二、布局形式方面：眼动打字布局应满足所使用眼动追踪设备分辨率要求，根据眼动追踪技术精度合理布局。在具体布局形式上，QWER 布局是用户最为熟悉的二维眼动布局方式，适用于二维显示应用场景；眼中心圆形布局更贴合人眼输入特点，适用于三维虚拟现实或混合现实应用场景。

三、反馈形式方面：视觉反馈仍然是软键盘反馈设计中最主要的设计方向，集中表现为交互控件颜色变化、交互空间图形变化等。多通道反馈形式能够有效提升交互绩效、缓解视觉通道压力，但应注意其他通道的反馈力度，不可干扰用户交互状态。

四、预测能力方面：具备字符或单词预测能力仍是大多数眼动打字软键盘的选择。将预测能力反馈至界面布局上可有效减少用户交互过程中的视线转移，将字符与单词预测合理融合至界面布局中，可有效提升眼动打字效率。

第三章 眼动打字软键盘设计方案

基于眼动交互特点与眼动交互软键盘的设计原则，本章将详细阐释所提出的三种眼动打字软键盘布局方案。

3.1 软键盘设计方案一

3.1.1 界面介绍

软键盘交互界面如图 3-1 所示。本界面为 Dwell-based 交互界面，主要面向虚拟现实等以头戴式显示器为媒介的三维交互场景。



图 3-1 方案一软键盘界面演示

在界面布局方面，考虑到以虚拟现实为代表的三维场景眼中心显示特点，主体界面将采用**圆形布局**。软键盘共涉及 29 个字符（26 个字母字符，3 个功能字符），各字符分布在**周边视觉区域**上；中央凹区域一方面用于显示输入字符，另一方面允许透明化显示虚拟环境或现实环境。

在命令触发方面，软键盘作为 Dwell-based 键盘采用**自定义注视时间**触发，用户可根据自身需求调节注视时间。此外，建立预测模型，基于上下文输入文本预测用户当前输入字符概率，**对于高概率字符的选择将缩短所需注视时间**。

在反馈形式方面，软键盘将注视过程以颜色表征的形式可视化，用户通过观察颜色填充扇形字符程度可直观察觉自己的注视时长。此外，根据用户视觉指向自动扩大指向按钮的扇形大小，根据**扇形大小与颜色填充**可得知当前的视线指向目标（详见图 3-2）。

在预测能力方面，软键盘的预测能力聚焦于两方面——**字符预测与单词预测**。字符预测表现在预测模型可根据用户上下文输入文本预测当前各字符的输入概率，概率最高的五个字符扇形半径增大以突出显示，且高概率的字符其注视时间将自动缩短；单词预测表现在预测模型可根据上下文输入文本预测当前指向最可能输入的单词，并以增强扇形的形式显示在当前指向字符的附近（详见图 3-2）。

3.1.2 使用流程

以输入文本“TEXT”为例，软键盘使用流程如图 3-2 所示。

(a) 输入字符 T。用户视觉搜索字符 T，字符 T 所在扇形被拉伸，且随着注视时间延长字符 T 逐渐被蓝色标识填充。各字符按钮在被指向时自动拉伸以扩大交互控件范围，减少因按键过小或视线抖动造成注视失败。

(b) 输入字符 T 之后。根据前文输入字符 T，系统将自动拉长预测概率高的五个字符以突出显示，这些字符的注视时间也被相应缩短。

(c) 输入字符 E。根据当前正在输入字符预测可能出现单词，最可能的单词被额外扇形突出显示。

(d) 直接输入预测单词 TEXT。完成字符 E 注视后，视线不要移开，继续注视字符 E，此时会有红色扇形进度条显示正在输入预测单词，等待红色进度条铺满扇形后将直接输入预测单词 TEXT。

(e) 输入 TEXT 成功。

(f) 若要输入 TEA 或 TEE，则在完成字符 E 注视后主动离开视线。此时软键盘将重新根据输入 TE 预测接下来的五个高概率字符并加长显示。用户可以继续选择输入 A 或 E，预测单词同样根据当前指向显示在字符附近。

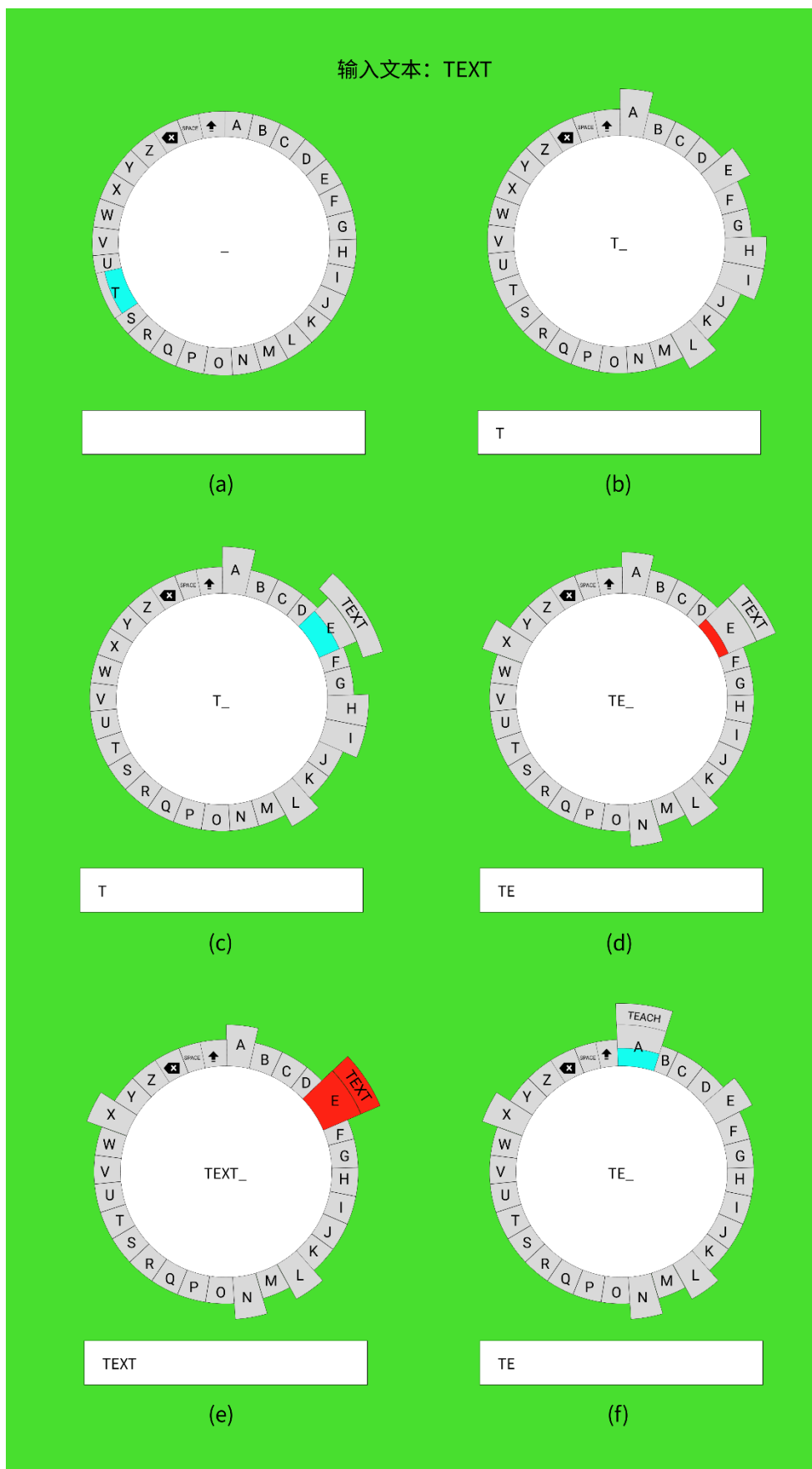


图 3-2 方案一软键盘使用流程

3.2 软键盘设计方案二

3.2.1 界面介绍

软键盘交互界面如图 3-2 所示。本界面为混合型软键盘交互界面，主要应用于显示器、渐冻症人显示设备等二维交互场景。



图 3-3 方案二软键盘界面演示

在界面布局方面，考虑到二维交互场景中 QWER 键盘布局对于用户熟悉度更高，本方案采用 QWER 布局形式。软键盘共包含 29 个交互字符（26 个字母字符与 3 个功能字符），软键盘最上方空白处显示预测单词，用户可通过注视预测单词完成单词输入。

在命令触发方面，本方案采用 **Dwell-based+Dwell-free** 的混合型触发方式。具体触发方式为：当用户想要输入某一单词时，首先通过注视单词第一个字符至字符变色，再通过平滑追踪依次扫视自己想要输入字符的视线轨迹，最终注视单词最后一个字符至字符变色，预测模型将根据**开始字符、结尾字符、视线轨迹**三方面分析预测出当前意图单词，并将单词显示在预测栏中。用户通过注视预测栏中的单词方案完成单词输入。**Dwell-based** 环节保证了用户单词输入意图与其他意图的准确分离，**Dwell-free** 环节加快了单词输入节奏，二者结合可实现输入稳

定性与输入效率的结合。本方案注视时间可根据用户自定义调整。（详见图 3-4）

在反馈形式方面，主要采用视觉反馈。被视线指向的字符按键将被字符框框起，并下方显示注视进度条。当注视条满时，字符被选中，如果该字符作为单词开始字符，则被标记为蓝色；如果该字符作为单词结束字符，则被标记为红色。用户注视预测框单词时，单词同样被框选，注视结束后单词将变蓝并被键入。（详见图 3-4）

在预测能力方面，本方案预测模型通过用户开始字符、结尾字符以及整个视线轨迹三方面完成单词预测。（详见图 3-4）

3.2.2 使用流程

以输入文本“this”为例，软键盘使用流程如图 3-4 所示。

(a) 选择开始字符 t。用户注视按键 t，此时按键 t 被框选且下方显示注视持续时间。

(b) 完成开始字符 t 的选择。用户注视持续一定时间后，按键 t 被作为开始字符所选中，此时变为蓝色。基于输入字符 t，预测栏将显示预测单词。

(c) 扫视文本 this。选择开始字符 t 后，用户依次扫视按键 h、按键 i、按键 s，并将实现停留在按键 s 上。根据开始字符与扫描路径，预测栏将更新显示预测单词。

(d) 选择结束字符 s。用户注视按键 s，此时按键 s 被框选且下方显示注视持续时间。

(e) 完成结束字符 s 的选择。用户注视持续一定时间后，按键 s 被作为结束字符所选中，此时变为黄色。根据开始字符、扫描路径、结束字符，预测栏将更新显示预测单词。

(f) 注视 this 单词。用户注视预测栏单词 this，此时单词被框选且下方显示注视持续时间。

(g) 完成 this 单词键入。用户注视持续一定时间后，单词 this 被键入，此时界面在下一刻将回归原始状态。

(h) 如果只想输入单个字符 t。在完成字符 t 的注视后，用户需再次注视预测栏中的单词 t 来完成单个字符键入。

输入文本：this



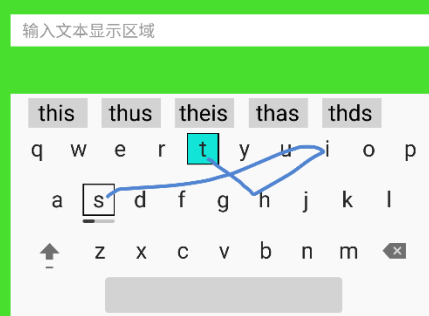
(a)



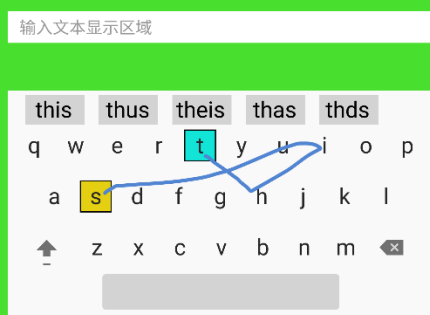
(b)



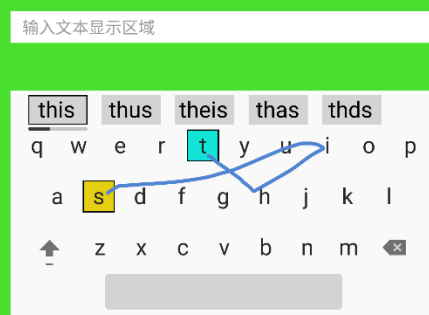
(c)



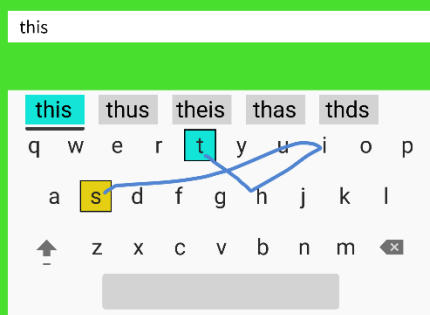
(d)



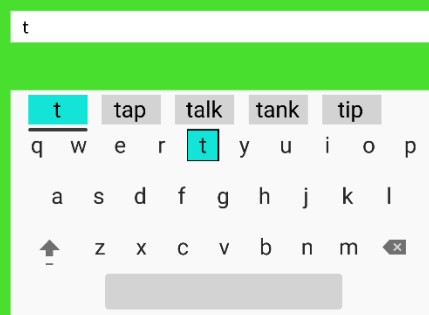
(e)



(f)



(g)



(h)

图 3-4 方案二软键盘使用流程

3.3 软键盘设计方案三

3.3.1 界面介绍

软键盘交互界面如图 3-5 所示。本界面为 Dwell-free 软键盘交互界面，主要面向虚拟现实等以头戴式显示器为媒介的三维交互场景。

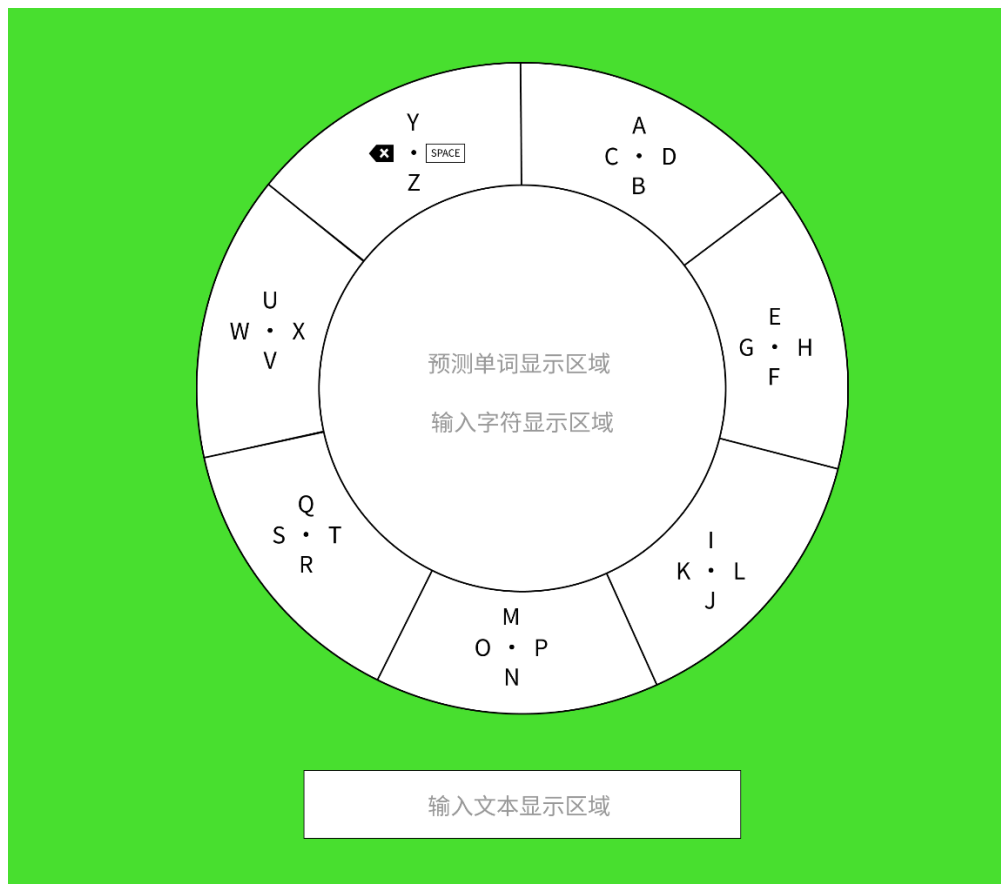


图 3-5 方案三软键盘界面演示

在界面布局方面，依旧采用符合人眼视觉搜索特点的圆形布局。界面共包含 28 个字符(26 个字母字符+2 个功能字符)，每 4 个字符为一字母簇，环形分布在周边视觉区域，4 个字符间采用上下左右分布。中央区域为输入字符显示区域，其上方为预测单词显示区域。

在命令触发方面，本方案采用 Dwell-free 方式，通过眼势动作完成字符键入。具体流程为：用户视线指向某一字符簇；字符簇弹出确认字符；用户做出视线沿目标字符方向移出扇形、再将视线移回原位的眼势动作；目标字符输入成功。该命令触发方式参考 EyeSwipe^[45]，如图 3-6 所示。

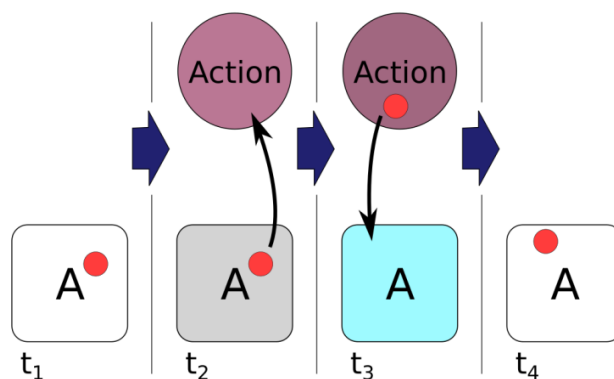


图 3-6 EyeSwipe^[45] 命令触发方式

在反馈方式方面，以视觉反馈为主。当用户视线指向某一字符簇时，字符簇所在扇形变为蓝色，且弹出红色圆圈的字符触发按键。当用户向目标字符做出“先移出后移入”的眼势动作后，完成目标字符选择，目标字符触发按键自动变暗。对于预测单词的选择同样如此（详见图 3-7）。

在预测能力方面，本方案采用根据已经输入的字符进行单词预测，并将预测单词显示在输入文本显示区域上部。

3.3.2 使用流程

以输入文本“TEXT”为例，软键盘使用流程如图 3-7 所示。

- (a) 输入字符 T 的第一阶段。视线指向 QRST 字符簇，字符簇所在扇形变蓝。
- (b) 输入字符 T 的第二阶段。字符簇展开触发按键，用户视线向按键 T 所在的右方移动。
- (c) 输入字符 T 的第三阶段。用户视线从右方返回中央区域，完成眼势动作，此时触发按键因被触发而变暗。
- (d) 输入字符 T 的最后阶段。字符 T 被输入，预测单词显示在上方。
- (e) 输入预测单词 TEXT。视线指向预测单词 TEXT，此时弹出选择触发按键。
- (f) 完成预测单词 TEXT 输入。做出“向上扫视再返回”的眼势动作，完成单词 TEXT 的输入。

输入文本：TEXT

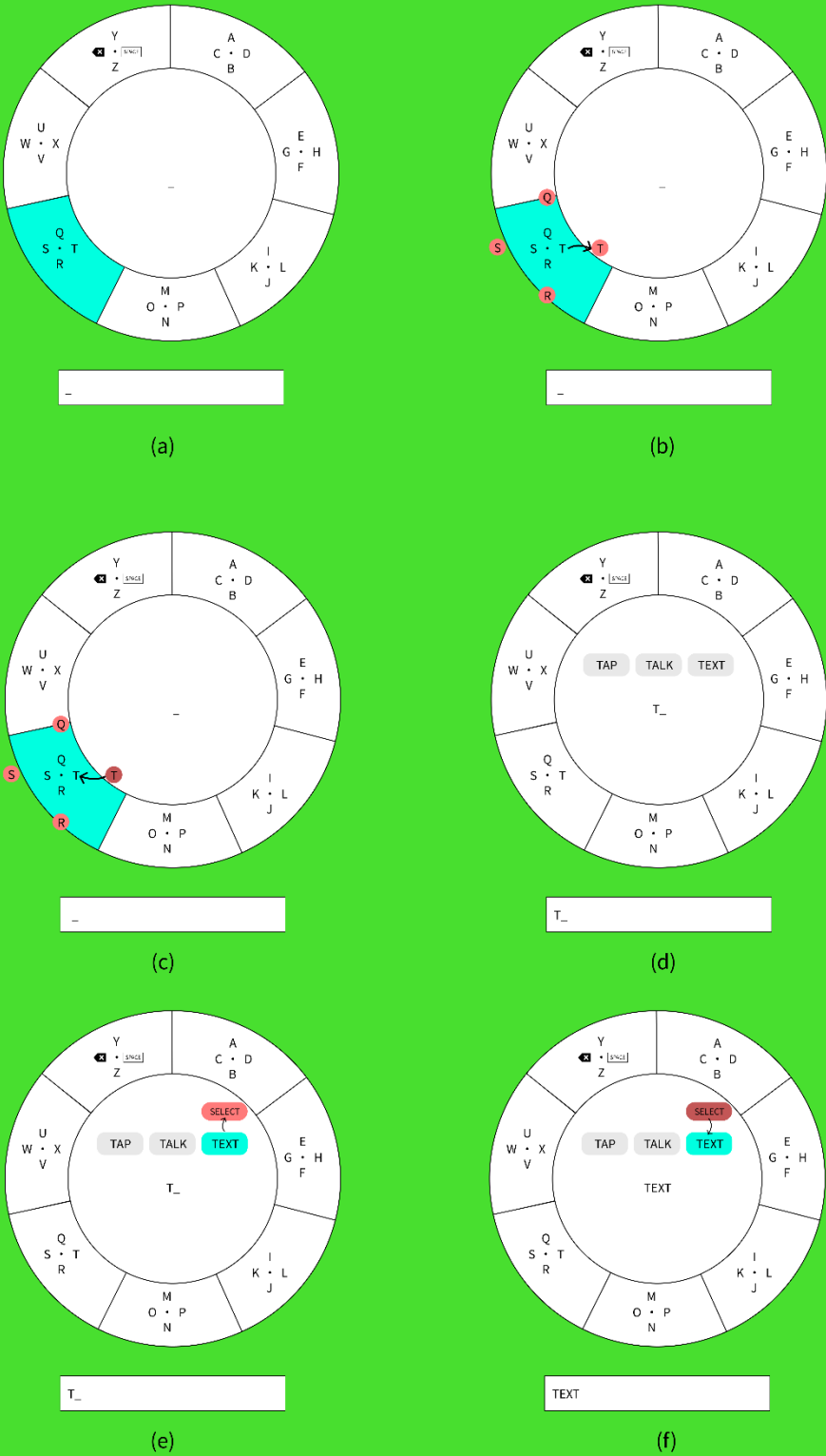


图 3-7 方案三软键盘使用流程

第四章 总结

人机交互设计强调计算机互动方式应契合人的生理与认知特点，基于眼动打字软键盘设计同样如此。聚焦人眼输入的不稳定性、聚焦特殊人群的生理障碍，眼动打字需要从用户角度、从交互情境角度进行交互界面设计与交互方式创新，通过提升眼动数据采集精度、改善更符合人行为的界面布局、建立人的眼动打字行为模型，进而构建无障碍的眼动打字系统。

本次课程报告从特殊人群的眼动打字需求出发，综述了面向行为的人眼运动特点与面向策略的软键盘布局设计案例，并给出了三个不成熟的眼动打字设计方案。人机交互无处不在，但面向自然高效人机交互的发展愿景，目前仍然有很长的路要走。

参考文献

- [1] 葛列众.工程心理学[M].上海:华东师范大学出版社,2017:163-164.
- [2] Jacob R, Karn K. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises[M]//The mind's eye. North-Holland, 2003: 573-605.
- [3] Papoutsaki A, Gokaslan A, Tompkin J, et al. The eye of the typer: a benchmark and analysis of gaze behavior during typing[C]//Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications. 2018: 1-9.
- [4] Klaib A, Alsrehin N, Melhem W, et al. Eye tracking algorithms, techniques, tools, and applications with an emphasis on machine learning and Internet of Things technologies[J]. Expert Systems with Applications, 2021, 166: 114037.
- [5] Mele M, Federici S. Gaze and eye-tracking solutions for psychological research[J]. Cognitive processing, 2012, 13(1): 261-265.
- [6] Bitkina O, Park J, Kim H. The ability of eye-tracking metrics to classify and predict the perceived driving workload[J]. International Journal of Industrial Ergonomics, 2021, 86: 103193.
- [7] Ryffel C, Muehlethaler C, Huber S, et al. Eye tracking as a debriefing tool in upset prevention and recovery training (UPRT) for general aviation pilots[J]. Ergonomics, 2019, 62(2): 319-329.
- [8] Pugliese R, Sala R, Regondi S, et al. Emerging technologies for management of patients with amyotrophic lateral sclerosis: from telehealth to assistive robotics and neural interfaces[J]. Journal of Neurology, 2022: 1-12.
- [9] Aoki H, Hansen J, Itoh K. Learning gaze typing: what are the obstacles and what progress to expect[J]. Universal Access in the Information Society, 2009, 8(4): 297-310.
- [10] Kar A, Corcoran P. A review and analysis of eye-gaze estimation systems, algorithms and performance evaluation methods in consumer platforms[J]. IEEE Access, 2017, 5: 16495-16519.
- [11] Jacob R. The use of eye movements in human-computer interaction techniques: what you look at is what you get[J]. ACM Transactions on Information Systems (TOIS), 1991, 9(2): 152-169.
- [12] Mele M, Federici S. A psychotechnological review on eye-tracking systems: towards user

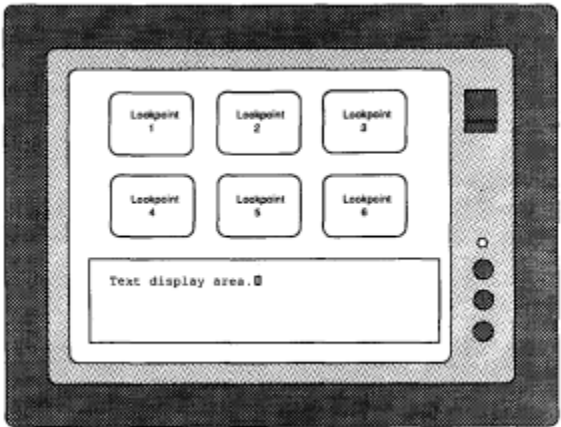
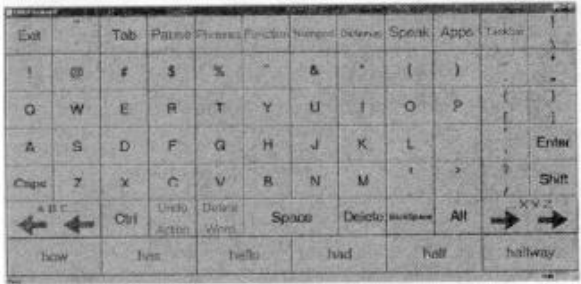

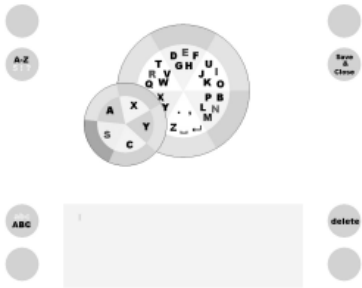
- experience[J]. *Disability and Rehabilitation: Assistive Technology*, 2012, 7(4): 261-281.
- [13] Just M, Carpenter P. A theory of reading: from eye fixations to comprehension[J]. *Psychological review*, 1980, 87(4): 329.
- [14] Hubel D. *Eye, brain, and vision*[M]. Scientific American Library/Scientific American Books, 1995.
- [15] Rayner K, Pollatsek A, Ashby J, et al. *Psychology of reading*[M]. Psychology Press, 2012.
- [16] Zander T, Gaertner M, Kothe C, et al. Combining eye gaze input with a brain–computer interface for touchless human–computer interaction[J]. *Intl. Journal of Human–Computer Interaction*, 2010, 27(1): 38-51.
- [17] Kristensson P, Vertanen K. The potential of dwell-free eye-typing for fast assistive gaze communication[C]//*Proceedings of the symposium on eye tracking research and applications*. 2012: 241-244.
- [18] Kowalczyk P, Sawicki D. Blink and wink detection as a control tool in multimodal interaction[J]. *Multimedia Tools and Applications*, 2019, 78(10): 13749-13765.
- [19] Blascheck T, Kurzhals K, Raschke M, et al. Visualization of eye tracking data: A taxonomy and survey[C]//*Computer Graphics Forum*. 2017, 36(8): 260-284.
- [20] Kar A, Corcoran P. A review and analysis of eye-gaze estimation systems, algorithms and performance evaluation methods in consumer platforms[J]. *IEEE Access*, 2017, 5: 16495-16519.
- [21] Cristina S, Camilleri K. Unobtrusive and pervasive video-based eye-gaze tracking[J]. *Image and Vision Computing*, 2018, 74: 21-40.
- [22] Jönsson E. *If looks could kill—an evaluation of eye tracking in computer games*[J]. Unpublished Master’s Thesis, Royal Institute of Technology (KTH), Stockholm, Sweden, 2005.
- [23] Helmert J, Pannasch S, Velichkovsky B. Influences of dwell time and cursor control on the performance in gaze driven typing[J]. *Journal of Eye Movement Research*, 2008, 2(4).
- [24] Wei-Gang C, Huang C, Hwang W. Automatic eye winks interpretation system for human-machine interface[J]. *EURASIP Journal on Image and Video Processing*, 2007, 2007: 1-9.
- [25] Rozado D, Moreno T, San A, et al. Controlling a smartphone using gaze gestures as the input mechanism[J]. *Human–Computer Interaction*, 2015, 30(1): 34-63.

- [26] 董占勋, 许若楠, 顾振宇. 眼动交互在大屏手机操作中的应用趋势[J]. 包装工程, 2015, 36(24): 57-60.
- [27] Chamberlain L. Eye tracking methodology theory and practice[J]. Qualitative Market Research: An International Journal, 2007.
- [28] Holmqvist K, Nyström M, Andersson R, et al. Eye tracking: A comprehensive guide to methods and measures[M]. OUP Oxford, 2011.
- [29] Møllénbach E, Hansen J, Lillholm M. Eye movements in gaze interaction[J]. Journal of Eye Movement Research, 2013, 6(2).
- [30] Majaranta P, Räihä K. Twenty years of eye typing: systems and design issues[C]//Proceedings of the 2002 symposium on Eye tracking research & applications. 2002: 15-22.
- [31] Hutchinson T, White K, Martin W, et al. Human-computer interaction using eye-gaze input[J]. IEEE Transactions on systems, man, and cybernetics, 1989, 19(6): 1527-1534.
- [32] Frey L, White K, Hutchison T. Eye-gaze word processing[J]. IEEE Transactions on systems, Man, and Cybernetics, 1990, 20(4): 944-950.
- [33] Lankford C. Effective eye-gaze input into windows[C]//Proceedings of the 2000 symposium on Eye tracking research & applications. 2000: 23-27.
- [34] Hansen J, Hansen D, Johansen A. Bringing gaze-based interaction back to basics[C]//HCI. 2001: 325-329.
- [35] Huckauf A, Urbina M. Gazing with pEYEs: towards a universal input for various applications [C]//Proceedings of the 2008 symposium on Eye tracking research & applications. 2008: 51-54.
- [36] Benligiray B, Topal C, Akinlar C. SliceType: fast gaze typing with a merging keyboard[J]. Journal on Multimodal User Interfaces, 2019, 13(4): 321-334.
- [37] Majaranta P, Ahola U, Špakov O. Fast gaze typing with an adjustable dwell time[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2009: 357-360.
- [38] Mott M, Williams S, Wobbrock J, et al. Improving dwell-based gaze typing with dynamic, cascading dwell times[C]//Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017: 2558-2570.

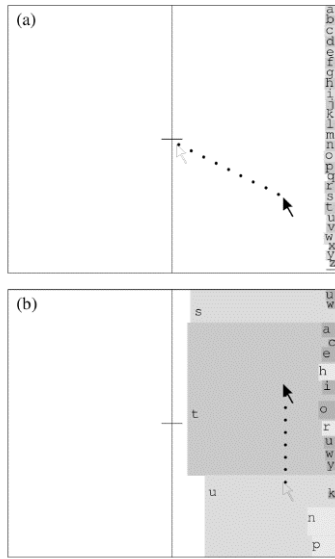
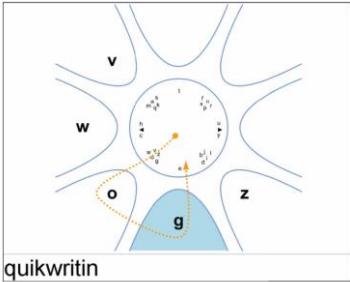
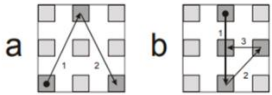
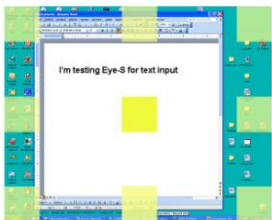
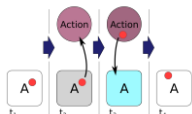
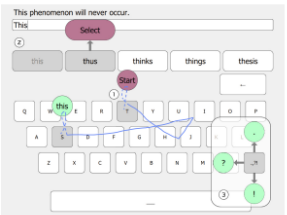
- [39] Pi J, Shi B. Probabilistic adjustment of dwell time for eye typing[C]//2017 10th International Conference on Human System Interactions (HSI). IEEE, 2017: 251-257.
- [40] Pi J, Koljonen P, Hu Y, et al. Dynamic bayesian adjustment of dwell time for faster eye typing[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2020, 28(10): 2315-2324.
- [41] Ward D, Blackwell A, MacKay D. Dasher—a data entry interface using continuous gestures and language models[C]//Proceedings of the 13th annual ACM symposium on User interface software and technology. 2000: 129-137.
- [42] Perlin K. Quikwriting: continuous stylus-based text entry[C]//Proceedings of the 11th annual ACM symposium on User interface software and technology. 1998: 215-216.
- [43] Bee N, André E. Writing with your eye: A dwell time free writing system adapted to the nature of human eye gaze[C]//International tutorial and research workshop on perception and interactive technologies for speech-based systems. Springer, Berlin, Heidelberg, 2008: 111-122.
- [44] Porta M, Turina M. Eye-S: a full-screen input modality for pure eye-based communication[C]//Proceedings of the 2008 symposium on Eye tracking research & applications. 2008: 27-34.
- [45] 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.
- [46] 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.
- [47] Liu Y, Lee B, McKeown M. Robust eye-based dwell-free typing[J]. International Journal of Human-Computer Interaction, 2016, 32(9): 682-694.
- [48] Scott M, Ashtiani B. BlinkWrite: efficient text entry using eye blinks[J]. Universal Access in the Information Society, 2011, 10(1): 69-80.
- [49] Xu W, Liang H, Zhao Y, et al. Ringtext: Dwell-free and hands-free text entry for mobile head-mounted displays using head motions[J]. IEEE transactions on visualization and computer graphics, 2019, 25(5): 1991-2001.

- [50] Lutz O, Venjakob A, Ruff S. SMOOVs: Towards calibration-free text entry by gaze using smooth pursuit movements[J]. *Journal of Eye Movement Research*, 2015, 8(1).
- [51] Zeng Z, Neuer E, Roetting M, et al. A One-Point Calibration Design for Hybrid Eye Typing Interface[J]. *International Journal of Human-Computer Interaction*, 2022: 1-14.
- [52] MacKenzie I, Zhang S, Soukoreff R. Text entry using soft keyboards[J]. *Behaviour & information technology*, 1999, 18(4): 235-244.
- [53] Al K, Mojahid M, Vigouroux N. Bigkey: A virtual keyboard for mobile devices[C]// *International Conference on Human-Computer Interaction*. Springer, Berlin, Heidelberg, 2009: 3-10.
- [54] Rajanna V, Russel M, Zhao J, et al. PressTapFlick: Exploring a gaze and foot-based multimodal approach to gaze typing[J]. *International Journal of Human-Computer Studies*, 2022, 161: 102787.
- [55] Niu Y, Tian J, Han Z, et al. Enhancing User Experience of Eye-Controlled Systems: Design Recommendations on the Optimal Size, Distance and Shape of Interactive Components from the Perspective of Peripheral Vision[J]. *International Journal of Environmental Research and Public Health*, 2022, 19(17): 10737.
- [56] Porta M, Dondi P, Pianetta A, et al. SPEye: A Calibration-Free Gaze-Driven Text Entry Technique Based on Smooth Pursuit[J]. *IEEE Transactions on Human-Machine Systems*, 2021, 52(2): 312-323.
- [57] Diaz-Tula A, Morimoto C. Augkey: Increasing foveal throughput in eye typing with augmented keys[C]// *Proceedings of the 2016 CHI conference on human factors in computing systems*. 2016: 3533-3544.
- [58] MacKenzie I, Zhang X. Eye typing using word and letter prediction and a fixation algorithm[C]// *Proceedings of the 2008 symposium on Eye tracking research & applications*. 2008: 55-58.

附录 I 眼动打字软键盘界面案例

序号	界面名/作者名	界面图
[32]	Frey	
[33]	Lankford	
[34]	GazeTalk	
[35]	pEYEs	

序号	界面名/作者名	界面图
[36]	SliceType	<p>(a) (b) (c) (d)</p>
[37]	Majaranta	
[38]	Mott	
[39]	Pi	
[40]	Pi	

序号	界面名/作者名	界面图
[41]	Dasher	
[43]	Quikwriting	
[44]	Eye-S	 <p data-bbox="821 1339 1061 1361">Figure 4: Possible eye sequences for the 'a' and 'b' letters</p>  <p data-bbox="853 1635 1029 1657">Figure 7: Explicit display of hotspots</p>
[45]	EyeSwipe	 <p data-bbox="614 1832 933 1888">Figure 1. To select the "A" key by reverse crossing, the user moves their gaze (red circle) to the key (time t_1) and the "Action" button is displayed (t_2). The user then looks at the "Action" button (t_3) and then back at the key (t_4). The "Action" button disappears and the "Action" is performed.</p>  <p data-bbox="965 1933 1284 1989">Figure 2. (1) To type the word "this", the user indicates the first ("T") and last ("S") characters by reverse crossing (dashed blue line), glancing through the vicinity of "H" and "I" (blue line). Candidates (2) and punctuation (3) can also be selected by reverse crossing.</p>

序号	界面名/作者名	界面图
[53]	BigKey	
[54]	PressTapFlick	
[56]	SPEye	
[57]	Augkey	