

人机交互技术期末大报告

项目名称：基于手势识别的 Do-Re-Mi 七音弹奏

Web 应用

摘要

本项目旨在构建一款基于 Web 端的人机交互应用，用户通过摄像头与手势识别技术控制七个音符（Do、Re、Mi、Fa、Sol、La、Ti）在页面上的圆形琴键排列，实现自由弹奏与旋律挑战双模式的交互体验。系统采用MediaPipe Hands 进行实时手势识别，结合 p5.js 实现图形界面与交互逻辑，同时加入节奏引导机制与按钮界面优化，增强用户参与感与趣味性。通过用户测试与反馈，验证了该系统在学习性、沉浸感和交互响应方面的良好表现。

I. 引言

1. 项目背景与目标

本项目旨在开发一个融合人机交互原理的音乐互动系统，鼓励用户通过手势识别技术直观地弹奏 Do-Re-Mi 七音音符，以提升音乐学习的参与度与趣味性。项目紧扣“自然交互”“视觉+听觉反馈”与“低门槛音乐体验”三大方向，结合课程目标，聚焦于 HCI 中多模态交互与沉浸体验的落地实践。

2. 中期报告回顾与项目演进

起初提出了“节奏解谜”的玩法概念，探索“自然派”（石碑、环境交互）与“琴键派”（抽象琴键界面）两个方向。在后期开发中，我们选定“琴键派”，并迭代为基于手势识别的 Web 音乐应用，扩展了圆形音符布局、旋律挑战模式和节奏引导机制等模块，体现了用户中心与设计迭代理念。

3. 报告结构概述

本报告将从需求分析与概念设计入手，依次介绍系统架构与实现、界面与交互策略、用户评估方法与结果、总结与未来展望，最后附上参考文献与相关素材。

II. 需求分析与概念设计

1. 问题空间探索与用户需求识别

本项目聚焦于传统音乐学习中对设备依赖强、交互方式单一的问题，提出通过手势识别替代传统输入设备，使音乐学习与娱乐变得更加直观与自然。结合“你不是用户”的设计原则，我们面向非专业用户与初学者开展交互设计。

用户画像

小红，9岁，小学生，对音乐感兴趣但未接触过钢琴，希望用简单方式玩音乐。

Leo，22岁，大学生，喜欢新颖 Web 交互，希望通过摄像头手势来参与轻量音乐活动。

使用场景

小红在家中使用平板打开网页，依次挥手触发音符完成“小星星”的弹奏任务。

2. 功能与非功能需求

功能需求：

七音手势识别（Do-Ti）

自由弹奏与旋律挑战模式切换

琴键高亮、音符反馈动画与音效

数据需求：无需数据存储，仅在客户端处理音符、位置、状态信息

环境需求：具备摄像头、运行现代浏览器、光线正常的使用环境

用户特性：适用于儿童、学生、音乐爱好者等初学者群体

3. 可用性与用户体验目标

易学：5 分钟内掌握基本操作

高效：音符识别响应低于 200ms

安全：提示误识别，提供重置按钮

愉悦：动画反馈、粒子飞出增强感官反馈

吸引人：环形音符、节奏指示提升沉浸感

有益：通过视觉与听觉双模反馈辅助音符学习

4. 概念模型与交互策略

系统以“琴键环”为界面隐喻，通过圆形音符布局将虚拟乐器转化为可直接操作的交互对象。核心交互方式为“操作型”（manipulation），即用户用食指位置在界面中物理触碰各个区域触发响应。

5. 认知方面考量

注意：使用音符高亮提示目标；节奏闪烁吸引视觉注意

感知：颜色编码与动画差异强化音高辨识

记忆：音符 Do-Ti 配色与标签降低记忆负担

学习：挑战模式逐步呈现目标旋律，支持探索式学习

6. 设计原则应用（Nielsen）

可见性：当前状态通过动画与文字反馈明确展示

反馈：每次弹奏伴随音效与视觉反馈

一致性：所有音符样式统一，按钮风格一致

可供性：琴键外形与粒子提示诱导交互意图

错误预防与恢复：误弹提示与音符重置机制

7. 伦理考量

系统使用MediaPipe 仅在本地识别摄像头图像，不上传或记录图像数据，保障用户隐私安全。界面提供摄像头使用说明。

III. 界面设计与原型实现

1. 设计流程

项目经历了从中期报告的平面原型，到后期 p5.js 高保真交互实现的全过程。

我们采用迭代方法，在每轮测试后优化琴键布局与界面反馈。

2. 具体界面设计

Web 页面以 640x480 画布居中展示，背景为柔和图片。

手势区域覆盖画布全部，使用 MediaPipe 进行 21 个关键点追踪。

图形元素包括圆形琴键、中心提示区、按钮组（切换模式、开始挑战等）。

3. GUI 元素

琴键风格：黑白配色基础上，赋予每个音符不同颜色标记，提高辨识度

反馈动画：触发音符会产生光效与

！

粒子飞出

多媒体整合：音符播放使用 p5.sound 音效库，与视觉反馈同步响应

4. 技术实现

前端框架：使用 p5.js 进行画布绘制与手势可视化

手势识别：基于 MediaPipe Hands 识别用户食指位置

交互逻辑：通过距离检测判断食指是否进入音符区域并触发响应

节奏挑战机制：设置旋律序列与定时器，逐步引导用户匹配目标旋律

IV. 人机交互设计

本系统在设计中充分考虑了用户、任务、系统三者之间的交互过程，围绕“低门槛、强反馈、高趣味”目标进行交互机制设计，主要体现在以下几个方面：

1 交互模型设计

采用直接操作模型（Direct Manipulation Model），用户通过实际手势动作直接控制音符触发，避免了复杂菜单与按钮操作；

系统识别食指位置，并实时反馈其在画布中的投影圆点，帮助用户建立清晰的

映射关系；

利用空间位置直观映射音高布局（左低右高），符合用户音乐认知习惯。

2 状态可视化反馈

可视化指引：通过食指实时位置圆点显示，降低用户迷茫感；

音符高亮反馈：当音符被成功触发时，圆形琴键变色并伴随粒子飞出动画，增强触觉缺失下的视觉反馈；

提示机制：系统启动时提供文字引导，首次操作后自动消隐，平衡新手引导与流畅体验。

3 容错与鲁棒性设计

通过合适的判定阈值设置，允许用户在小范围偏差内依然判定为成功触发，提升系统容忍度；

食指追踪稳定性优化，防止轻微抖动导致误判；

提供退出自由，无需复杂重置流程，鼓励自由探索式演奏。

4 学习与适应性设计

系统整体无学习成本，用户可在 1-2 分钟内掌握全部交互逻辑；

通过简洁的目标（手指触碰圆圈产生音符）快速建立心理模型；

后续可扩展引入节奏训练与挑战任务，逐步提升用户掌握深度。

5 情感化设计元素

背景图像柔和舒适，营造轻松愉悦的交互氛围；

音符与动画同步，提升成就感；

粒子音符 🎵 视觉特效，增加趣味性与吸引力；

适当使用拟物化符号，降低认知负荷。

V. 结论与反思

本项目以“基于自然手势的人机互动”为出发点，成功构建出一款融合自由弹奏与引导学习的 Web 端音乐应用。设计过程中始终秉持“以用户为中心”，通过迭代优化实现了功能性与趣味性的统一。项目不仅增强了对交互设计方法与评估流程的理解，也加深了对手势识别与 Web 多媒体编程技术的掌握。

VI. 参考文献

人机交互课程 PPT：第 0-15 讲（课程资料）

MediaPipe 官方文档：<https://mediapipe.dev/>

p5.js Reference：<https://p5js.org/reference/>**VII. 附录**

项目中期报告（hci_report.pdf）

系统源代码及说明文档（<https://github.com/devil-crash/Music-of-Index-Finger>）