**清华大学硕士学位论文**

**基于动态难度调节（自主调节）的视唱游戏设计研究与实现**

院 系： 深圳国际研究生院

专 业： 电子信息（互动媒体设计与技术）

姓 名： 张嘉睿

学 号：2020214447

导 师： 倪士光

2022年 10 月 22 日

目 录

第1章 绪论 3

1.1 研究背景 3

1.2 研究内容……………………………………………………………………4

1.3 研究意义 5

第2章 研究概况 6

2.1 视唱概述 8

2.2 ICT视唱练耳的应用现状 8

2.3 现阶段ICT实现方案的缺陷 10

2.4 DDA研究现状 11

2.5 OLMs研究现状 16

2.6 曲目难度评估方法研究现状 16

第3章 研究内容与方法 18

3.1 研究假设 18

3.2 研究对象和框架思路 18

3.3 研究目标 19

3.4 研究方法 20

3.5 研究内容 20

3.6 研究创新点 27

参考文献 27

# 第1章 绪论

## 研究背景和意义

### 研究背景

音乐教育是落实“素质教育”的重要载体。十九大明确指出，要重视音乐教育的发展和审美能力的培养。音乐不仅是提高学生文化素质水平、审美水平的重要手段，也对儿童的心理和智力发展有积极的影响，据[1][2],音乐教育对认知功能的提升有规划、工作记忆、语言智商和抑制功能。而视唱练耳旨在培养学生对音高、音长、音色的感知能力。学生能够注意和区分简单的节奏和旋律模式，以及不同节奏之间的关系。通过系统地接触音乐，能够培养对美的感知、对优质音乐的接受和评价，为良好音乐品位的培养奠定基础。

由于大众传媒的发达，人们现在可以以相当低的成本获取音乐甚至创造音乐，MIDI、录音、声音处理的成本越来越低。另外，日常环境中，音乐充斥着几乎每一个角落，人们不可控制地在商场等环境中接收音乐，且在私人场景中使用，以创造特定的情绪状态，或缓和兴奋水平[3]。

音乐教育面临的困难主要来源于几个不同的方面。首先是音乐课程设计本身的问题。音乐课程通常有非常明确和统一的书面目标，很少考虑个人的音乐经验；课程计划过于死板，几乎没有灵活的调整空间；对学生没有吸引力。

其次，师资问题。视唱练耳教学对教师的专业能力和教学能力都提出了较高的要求。但在我国，应试教育的风行导致了音乐教育的边缘化，师资质量难以得到保证。而在师资匮乏的乡村地区，推进音乐教育也是促进教育公平的重要部分。

2020年，在全球范围内COVID-19大流行的背景下，教育部发文表明，实施在线教育是推进教育公平的战略选择。而如何破解优质数字教育资源匮乏、教师信息技术应用能力不足的问题，成为本文研究的出发点。因此，我们提出该课题：基于动态难度调整的视唱练耳游戏研究与实现。研究聚焦如何推进基于DDA的个性化的游戏进程，有效提升学生的学习动机和学习效果。该应用部署于便携设备，支持学生的独立学习并提供及时反馈，以提供低成本、高质量的音乐学习体验。

### 研究意义

个性化教学是提高教学质量、培养创新型人才的关键，但在传统的教学环境下，由于教学资源和教师实力的限制，很难实现每个学生一对一的个性化教学[8]。随着网络教学的兴起，传统教学方法在时间和空间上的局限性得到了解决，基于网络的教学平台构建了虚拟化的教学情景和教学资源，使教学行为不再局限于课堂和课堂。学生可以通过无限时间和地点的在线学习平台自由学习，平台为私人教师等学习者提供各种服务，提供一对一的指导和帮助[9]。然而，仅仅为学习者供学习资源和学习环境并不能真正实现个性化教学，还需要学习平台需要与学习者积极互动，了解他们的个性特征，通过模仿教师的教学风格为不同的学习者提供有针对性的教学服务。智能教学系统利用人工智能技术模拟和学习人类教师的教学风格，根据不同学习者的需要为他们提供个性化的学习路径和学习建议，通过语言交流、行为分析和学习过程中的模拟来了解学习者的学习习惯和偏好，动态地引导学习者完成知识的掌握[10]。

综上所述，为音乐教育的游戏化设计做创新探索，尝试音乐教育在游戏理论与思想方向上的突破，找到学习和娱乐的结合点，并在此基础上挖掘学生的学习兴趣和提升学习效果，成为本研究的学术宗旨。

## 1.2中外研究现状综述

### 1.2.1视唱练耳游戏研究现状

视唱是基础音乐理论教育中理论知识和实践技能最重要的技能之一。视唱是音乐教学的一门基础性学科，在中国，从义务教育到专业音乐学院，再到综合性大学音乐系，都有与视唱相关的课程。学习和实践它使学生能够识别和理解符号和音乐符号的意义，并将它们与它们所代表的声音和节奏联系起来[11]。传统的视唱教学方法通常是由教师演奏一种乐器，然后监督学生重复、模拟或通过手指敲击练习节奏[12]。传统教学一般以纸质手册为基础，学生据此学习理论，并根据课后练习开展练习。随着信息时代的来临，视唱练耳的视频首先弥补了纸质手册在感官体验上的缺陷。

视唱是音乐教学中一门技术性很强的基础性学科。视唱教学的内容既包括基础音乐理论的学习，又包括理论知识技能的应用，因而具有综合性学科的特点视唱教学的内涵是通过对音乐节奏、音高、音调感觉和音乐阅读能力的训练，培养学生对音乐的感知、理解、欣赏和创造能力，使学生掌握正确的音高、节奏和表达音乐情感的能力，从而更容易地理解音乐形象，感知音乐的风格、类型和主题，为今后的音乐学习打下坚实的基础。这意味着视唱教学不仅要提高学生的视唱能力，还要培养他们的音乐能力。研究表明，视唱音乐能力 的培养是通过提高学生的节奏感、旋律感、强度感、和声感来实现的。

(1) 节奏感是学习者感知和再现音乐作品中节奏变化的能力。 节奏感的培养和训练提高了学生在音乐表演中调节呼 吸和节奏的能力，是培养学生感知、理解和创作音乐 能力的基本工具之一。

(2) 旋律是音乐作品中关键、节奏、节拍等音乐元素的有机结合，蕴含着丰富的艺术形象。准确地感知和表达音乐作品的旋律是一个人综合音乐能力的关键指标。只有通过准确地感知一部音乐作品的旋律，才能对音乐作为一个整体有一个系统而全面的理解，才能用最好的演唱方法来表达音乐中的艺术形象。

(3)力量是音乐作品表达情感的主要手段，掌握音乐作品的力 量结构，才能恰当地表达音乐的情感，情感的培养是对音乐情感表达能力的锻炼，可以增强学生歌唱的感染力和吸引力。

(4) 和声是两个或两个以上音调根据特定的规则和标准形成的固定的声学组合。和声反映了音乐在风格和类型上的一些特点，对音乐主题的确立具有重要意义，培养学生的和声意识有利于提高学生的层次感和音乐表现的三维性，增强学生的协调与合作能力。

随着ICT(Information and Communication Technology)的发展,ICT对音乐理论教育的重要性和作用被证实[13]。

1.2.2 视唱教学理论研究现状

1.2.2 ICT视唱练耳应用现状

（1）多模态视唱练耳应用程序：mySolfeggio[14]

该研究改善了传统视唱练耳学习的反馈匮乏的问题，并使教师可以在教学过程中查看学生的进度。应用程序将视觉、听觉和触觉融入到训练过程中，通过自动声音识别(ASR)、节奏识别、光学音乐识别(OMR)，加强学生对音高、节奏和和弦的识别能力。此外，该程序接入了增强现实技术(AR),用户可以将摄像机对准乐谱，并使用AR工具识别乐谱图像，从而将这些乐谱加载到游戏中。

使用该应用程序的学生在记谱准确性、音程和节奏之间比例的准确性方面得分高于仅使用传统方式练习视唱练耳的学生。但限于研究者的技术能力，节奏识别有延迟，而节奏识别在商业游戏中的延迟已几乎被消除。

（2）游戏化的视唱练耳平台：Troubadour[15]

该应用的主要形式为播放音程序列给学生，将第一个音符显示在五线谱中，学生的任务是识别听到的音符，并将它们输入到显示的乐谱中。

研究着重于利用游戏化元素优化用户体验，提升他们的投入，从而提升他们的学习效果，游戏化元素为徽章和排行榜。另外，游戏设置了练习的四个基本难度级别，难度与音乐学院的学期相对应。而在每组训练中，随着游戏的进行，根据学生的表现，听写的难度会增加，首先，连续的音符数量增加；其次，序列中出现更多的半音、三全音、不和谐音。音符是根据预先收集的音程分布自动生成的，复杂度主要取决于音符序列的长度、间隔的大小和间隔出现的频率。但文章没有对难度调节提供数据和学术支持。

（3）商业视唱练耳应用：Earpeggio

目前App Store中发现的最全面的听力训练应用，包括音程比较、音程识别、和弦识别、和弦转位、和弦进行等。应用对这些训练模块进行的分类，以供学生自行选择。但应用没有课程组件，也没有难度的渐进，学生需要教师的帮助才能进步。但Earpeggio提供了统计数据接口供教师查看，老师可以根据学生的能力调整进度。尽管如此，题目的设计有大量重复的内容，对题目的难度评价也模糊缺乏理论依据。

（4）基于提示衰退的视唱练耳训练应用 [16]

该应用主要采取的策略为提示衰退，即先在训练中提供色彩的提示，再逐渐使提示衰退吸引用户注意，并通过色彩刺激学习者，帮助学习者建立视觉上的心理模型。具体实现方式是，先呈现键盘钢琴的画面，将播放的音频信息利用颜色映射到钢琴键盘上，再逐渐将颜色褪去。该方式能够有效地提升听觉训练结果。

（5）基于虚拟现实的视唱练耳训练系统[17]

该研究主要设计了一种新型的虚拟现实练耳系统，用户调研显示，他们对系统的沉浸度很高，但虚拟现实提供的立体平移声效并不会对音程识别带来更好的训练效果，且研究者对UNITY音效的开发能力有限，不具备实现360°立体声模拟的能力。

（6）针对初级音乐学生的节奏游戏

这款游戏由几个独特的节奏小游戏组成，玩家必须配合歌曲的节奏及时玩游戏，以获得高分为目标。每个游戏时长1-3分钟，具体时长取决于游戏歌曲的长度。玩家按下节奏与游戏互动的方式因歌而异。游戏可能会要求玩家跟着节奏打拍子，对音乐或视觉线索做出反应，或者重复一种节奏模式。本次研究选择的三款游戏涵盖了所有列出的游戏类型：保持节拍(Air Rally)、回应提示(Hole in One)、重复一个模式(Rhythm Tweezers)。

保持节拍(Air Rally)：

卡通猫狗一边飞飞机一边打羽毛球。玩家扮演狗的角色，必须轻击将小鸟送回给猫。猫在数到1和3时击中小鸟，而玩家在数到2和4时负责将小鸟打回去。偶尔，猫会提示玩家等到数到3再打小鸟，然后再回到他们通常的节拍2和4。

回应提示(Hole in One):

玩家角色和一只猴子和一只山魈打高尔夫球。在投掷高尔夫球时，猴子会给玩家一根两拍的球杆，在球杆的第二拍之后，玩家必须拍一拍才能击球。一共有四拍球杆，球员必须在球杆的第四拍击球。

节奏镊子（Rhythm Tweezers）：

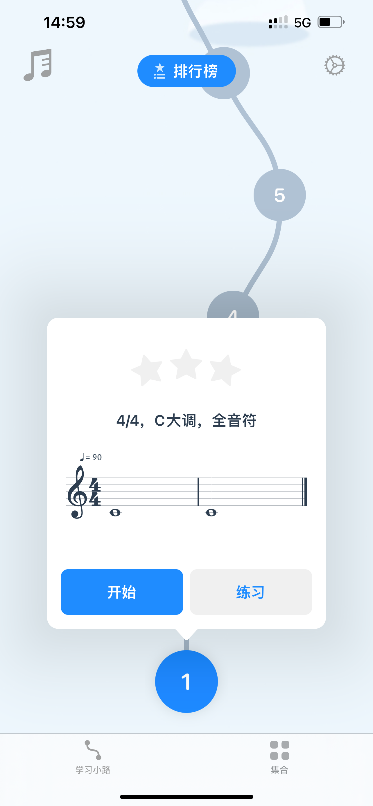
玩家需要控制一对镊子，把各种蔬菜的毛拔下来。这些毛发以4拍的节奏模式生长，玩家必须复制这个模式才能成功拔下这些毛发。

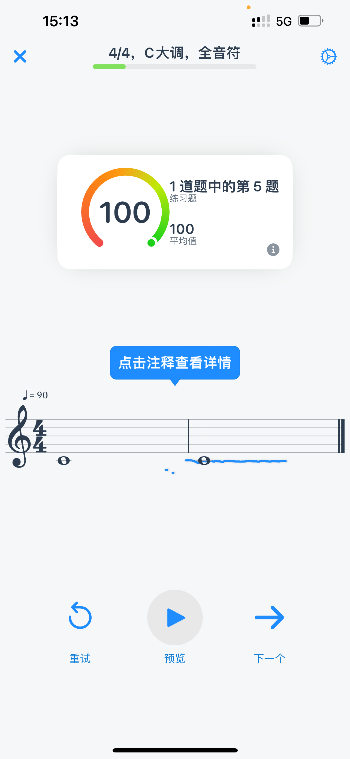
（7）新曲视唱

这款应用的视唱训练具有简单的UI反馈，用户哼唱的音高以音符的方式被记录在曲谱内。偏差较小的音符将呈现绿色，偏差较大的呈现红色。

曲目选择方式：仅有的选项为下方的“下一题”，且同难度等级的曲目数量过大，同质化严重，交互性较差。

游戏化设计：无

 自适应难度系统：无

（8）Ella

这款应用的市场训练也具有简单的UI反馈，用户哼唱的音高以曲线的形式记录在曲谱上，并给予节奏和音高上的打分，但精确度不高。

曲目选择方式：必须在通过前置任务的情况下，才能够解锁接下来的任务。

游戏化设计：排行榜

自适应难度系统：无。

|  |  |  |  |
| --- | --- | --- | --- |
| 视唱练耳训练 | 优点 | 缺点 | 是否有效果提升 |
| mySolfeggio | 教师监督  多模态 | 技术限制导致的节奏偏差 | 是 |
| Troubadour | 游戏化更成熟 | 难度曲线的调整缺乏依据和颗粒度 | 是 |
| Earpeggio | 内容全面  教师监督 | 大量重复 缺乏难度依据 | 是 |
| 提示衰退 | 设计创新 | 任务单一、缺乏编排和反馈 | 是 |
| 虚拟现实 | 沉浸度高 | 开发能力限制  音频技术能力限制 | 否 |
| 节奏天堂 | 具有更好的娱乐性，场景针对儿童进行了专门的设计。 | 训练模式单一且单局非常短暂 | 有 |

总的来说：

(1)在上述方案中，应用的难度设计存在缺失或缺陷。唯一有课程编排的Troubadour，其音符序列生成的难度依据，无论是学术的还是统计数据的，都没有在文章中给出。早期有研究对不同音程实例的难度进行了排序，但仅是从宏观人群的正确率进行统计。然而，个体在音乐听力任务的子任务能力水平存在差异[18]，具体如音高、节奏、音色的感知，音乐工作记忆水平、对不同音符排列模式的熟悉程度，均存在不同程度的差异。单一的难度评价体系对个体的适用性较差。

（2）游戏化方案依赖外部动机。现有的游戏化实现方式主要以排行榜、徽章体现。然而，赢得游戏的目标是一种外部动机，与用户实际接受的任务是分离的，而由于任务的视觉呈现方式和训练模式基本与纸质训练一致，只提供了基本的反馈功能，任务本身的内在动机几乎没有被加强。据研究[19]，与任务脱离联系的外部激励因素存在一些负面影响，比如用户无法专注于激励因素之外的内容。而让用户专注于任务的反思对学习来说是必要的，因此需要将正反馈与任务本身更紧密地结合。

（3）忽视了学生在安排学习内容上的自主性。上述研究大多提供了教师监督学生和安排学习任务的功能，但没有一个应用提供了学生自我管理的方法。研究表明，在游戏任务执行时，自主选择任务的用户内在动机水平更高[20]。

（4）歌唱学习评估标准在设计和实现方面缺乏依据。目前对难度推荐系统的设计和实现方法研究较少，视觉唱歌学习系统在个性化和智能化方面还不够完善[21]。其主要表现是学习资源管理和推动缺乏针对性，学习者能力水平评价依赖于人工判断，不能完全自动化和智能化。只考虑模板和表现之间差异的评估方法具有很大的错误[22]。此外，与曲库缺失相同的问题是，成绩的难易程度没有标准化的分类，它依赖于专家的主观判断。视唱教学评价中的问题只能来自专家数据库或教材，问题的个人偏好和专业水平会影响试题难度的平衡，不能保证评价的公正性、准确性和客观性[23]。因此，实现自动评分和提供基于评分难度差异的问题对于目前的视唱系统来说是一个挑战[24]。

### 1.2.2 DDA研究现状

动态难度调整是一种自适应地改变游戏属性以使游戏更容易或更难的技术，这种技术帮助开发者在游戏中实现良好平衡和优化玩家体验。为了保持玩家的专注度和参与度，游戏的挑战应取决于玩家的技能水平[25][26]。因此，游戏的难度必须得到控制，即对玩家来说不能太难也不能太简单。所有的DDA技术都有一个共同的步骤，即测量玩家在游戏中某一时刻面临的难度。

由于视唱练耳任务基于感官智能，玩家输入的信息仅限技能水平和使用意愿，基于人格预测和情感测量的DDA技术不作讨论。

由于视唱练耳游戏不存在博弈，因此基于决策过程或拟人的DDA也不作讨论。

2.4.1 Hamlet系统

应用于游戏《半条命》的动态难度调整系统，这个库包含以下函数：

1. 根据预先定义的参数监控游戏数据
2. 确定调整行动和策略
3. 执行调整行动和策略
4. 显示数据和系统控制设置
5. 生成游戏内的会话跟踪

当玩家在游戏中行动时，此系统将会使用统计参数监控传入的游戏数据，并根据这些数据预估玩家的未来状态。当预测到可能使玩家不悦且可以避免的状态，系统会调整游戏设置。

系统的关键步骤是，1.评估何时需要调整。2.确定应该进行哪些更改。3.尽可能地进行无缝更改。

1. 系统以玩家收到的伤害为准，统计游戏中发生的各种伤害事件发生的概率和伤害量，计算出给定时间点上总损害的预期。以玩家实际受到的伤害和预期收到的伤害差值为准，计算出玩家需要帮助的时机。
2. 调整行动主要有：改变已经出生的敌人命中率或伤害或生命值；改变未出生敌人的种类、出生时机顺序、生命值等。
3. 将调整放置在距离玩家更远的位置，当然这将影响调整的有效性。

2.4.2 基于玩家状态转移概率有向图的DDA[27]

该方法的目标是通过调整难度最大化用户粘性，奖励函数为用户总游玩回合数R，即玩家在整个游戏中所玩的预期总回合数。目前应用于EA的部分游戏。 随着游戏进程的推进，玩家从一个状态转移到另一个状态，而状态之间的转移概率取决于这些状态下的游戏难度。

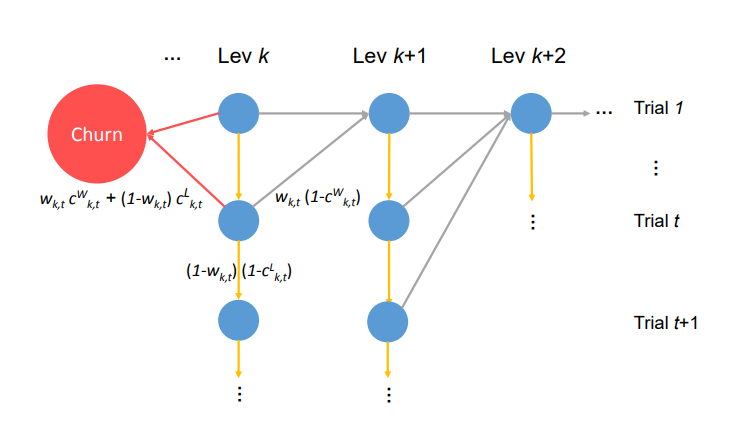


图2.1 玩家游戏状态转移概率图

图中，Lev 代表关卡编号，Trial代表关卡通关失败后的尝试次数。表示玩家在第次尝试进入第关，玩家到下一状态的转移概率依据通关失败次数和目前所处的状态计算。

是玩家获胜和升级的概率，是玩家失败后再尝试的概率。如果玩家流失，将不会增加。通过动态规划求解最优的，以最大化。

状态转换中采取流失率和作为已知参数，但只有和是不够的，忽略了玩家的其他个人特质。为了提高计算流失率的准确性，利用了流失率预测研究，也就是对单个玩家的特征生成复杂的预测模型来修正，计划在每个状态下预先计算可能的流失率。

采用此方法的小组在用户粘性指标方面有明显提高，可以拓展到其他的各种游戏类型。

2.4.3 张量分解方法预测游戏中的玩家技能使用的表现[28][29]

使用数据驱动建模方法明确预测玩家表现的变化，结合玩家之前的信息，预测构建挑战。研究者用张量表示玩家数据，是用户，是任务，是表现记录的时间，并根据上述因子进行关键信息提取以提升计算效率，此玩家模型预测玩家在未来技能事件中的表现。该方法相比非张量分解方法更优。

2.4.4 动态脚本[30]

游戏针对不同类型的敌人设计了不同的规则，这些规则是利用特定领域的信息人工设计的。创建敌人将根据他们的类型从规则库中抽取。选择规则的概率取决于分配给规则的权重。规则库通过修改值进行调整，反映相关规则的失败或成功率。

在这种方法中，学习是循序渐进的。在玩家完成一个遭遇战后，规则权重将根据其对结果的影响进行处理。玩家在某一方面的成功导致权重上升，而失败则会导致权重下降。其余的规则也将对这个调整进行适应，以保持所有规则的总权重不变。

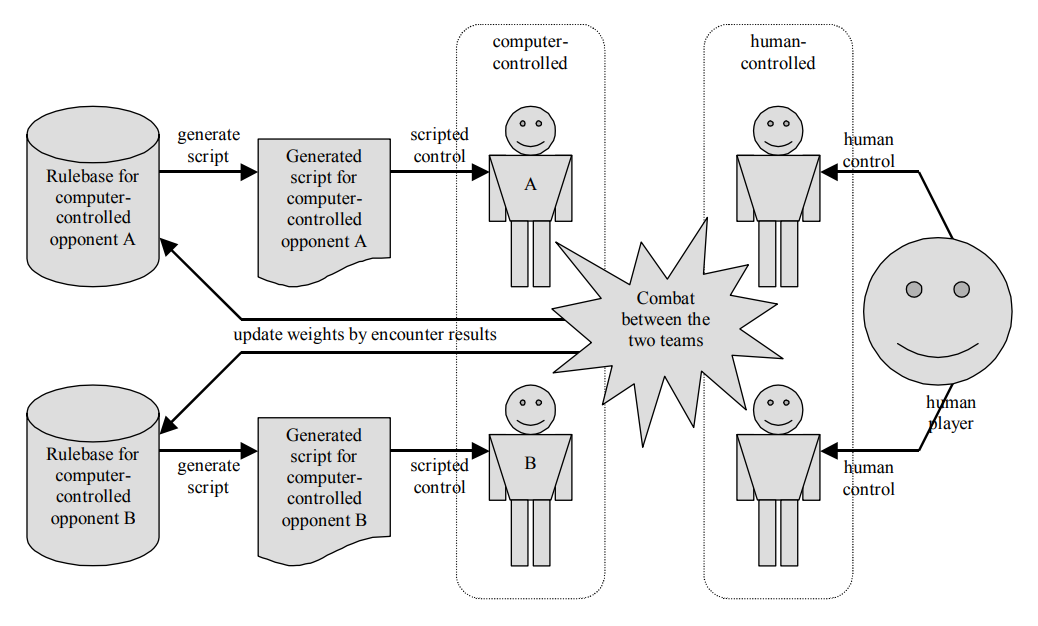


图2.2 动态脚本生成步骤图

为了保证敌人的表现适应玩家，另外设置了三个增强元素：

1. 高水平惩罚：为了让AI的表现水平配合玩家水平，使玩家感知的游戏难度适中，AI采取更中庸的策略取代最优策略，系统会调整权重，奖励AI的平庸值，惩罚过优值。
2. 权重裁剪：权重的最大值代表了一个策略能达到的最高优化水平。如果缩减最大值，可以限制权重的增长，导致许多策略都不是最优的，以提供游戏更好的平衡性。
3. 顶剔除：允许权重增长到最大值，但生成的脚本不会选择权重超过最大值的规则。如果AI频繁胜利，那么优势的策略将会被暂时屏蔽。

2.4.5 基于遗传算法的DDA关卡自动生成[31]

该方法先评估玩家表现，生成玩家的代理，再基于这种表现输出关卡。生成与玩家适应度相似的代理的步骤是1.先在随机生成的关卡集合中衡量玩家和所有可用代理的表现，并选择最能模拟玩家水平的代理。2.为关卡生成器随机创造一个初始参数集。3.为种群的每个代理生成一个关卡，并记录这些代理在这些级别上的表现。4.使用这些代理表现记录值计算种群中每个代理在这些关卡中的适应程度。5.基于每个代理的适应度值，使用遗传算法进化种群。6.达到期望的世代数就停止，否则从第3步再开始。

上述公式(1)(2)为难度计算函数，为难度值，为得分率，取值范围为,为完美通关时理论上的分数。采用哪种难度度量取决于关卡的任务。

公式(3)定义了关卡的适应度，为生成关卡的代号，为观测到的通关率，是任务目标的解决率。这个公式使我们能够将某个特定的代理所生成关卡的期望难度定义为一个百分比值(例如，如果希望代理平均每次通关进度为50%，那么)。

一旦计算出当前种群中每个参数的适应度值，就会使用遗传算法来进化种群并创造下一代。

该方法也被证明可以增加用户粘性。

2.4.5 动态时间正则化 (DTW)

动态时间正则化 (DTW)算法是一种基于成绩难度的资源推荐算法，在学习过程中将成绩按照适当的难度水平推送给学生，可以提高学习效率和评价效果，DTW 算法目前更常用于差异比较，使用距离来表示相似性。不同的研究根据评分目的和焦点选择不同类型的特征或组合不同类型的特征，评分效果的差异都 取决于所选择的特征。

### OLMs研究现状

许多现有的DDA教育游戏模型，往往基于学习者的表现，生成为每种结构特殊优化的推荐系统。这种系统自驱的适应性系统已经在视唱练耳游戏中实现。

人们正在积极研究开发引导学习者适应性特征的潜在模型，努力与学习者建立联系，提高系统透明度。由此研发的OLMs方法[32] ，通过记录学习者数据生成用户特征，并以此为凭据提出最佳学习路径[33],面向用户生成可调节的难度表盘，这些模型的使用已经导致学生在各个领域和系统中的学习得到改善[34]。

Lynnette[35]是一个使用 OLMs 来支持学习和自我调节的著名例子，该游戏为学习代数的应用，游戏中呈现了一个界面，显示学生在每个问题类型上的表现，同时也提供了关于问题选择的即时反馈; 学生在选择他们需要改进的问题类型时会收到正面反馈，而在选择他们已经掌握的问题类型时会收到负面反馈。Long和Aleven[36]的一项随机实验结果显示，与系统完全控制问题选择时相比，当学生被允许选择要做的问题类型并收到关于他们选择的反馈时，他们的学习效果更好。

### 1.2.4曲目难度评估方法研究现状

在音乐教学中掌握乐谱的难易程度，不仅有助于学生根据自己的水平选择合适的乐谱作为学习材料，而且有助于提高教学评价的公正性。虽然没有明确的标准来分类乐谱，但是在乐谱网站、音乐等级考试和教科书的布局上有乐谱难度的分类[37]。例如，乐谱下载网站根据大致的音高,音阶,和声和节奏将分数分为10个分级[38]。

根据音阶、和声、演奏方法等因素，将英国音乐演奏评分系统分为 8 个等级，为评价音乐演奏和理解提供了一个评分系统。就美国NYSSMA而言，根据音阶和八度音阶、音阶、演奏速度等，分为6个等级标准，这些标准是根据音乐的难度主观地分类出来的。视唱和听评分系统(SSMA)的教学大纲也根据乐谱的音高和节奏的复杂程度，将教学内容从容易到难度进行分类[39]目前对钢琴乐谱难度等级识别的研究主要是基于现有教科书或权威机构给出的具有明确难度等级标签的乐谱，并利用线性回归支持向量机等方法建立难度识别模型，试图利用该模型预测大量未标记乐谱的难度等级，以实现难度等级对乐谱进行自动分类[40]。旋律复杂性已经被研究了很长时间，它已经被用作研究的主要指标，如唤醒艺术刺激的潜力和享乐价值(偏好)音乐流行因素的分析[41]。这些研究说明了旋律的复杂性作为一种诊断工具的潜力，以确定何时出现旋律偶然性，或者旋律的哪些方面(音高组合、层次结构、顺序等)对于分析旋律结构(相似度评级、旋律曲目等)更重要[42][43][44][45][46]。

Eugene Narmour首先提出旋律复杂性的概念，他认为内隐实现模型在描述原始民歌的旋律复杂性方面表现良好，但缺乏对听众音乐和文化背景的考虑。为此，Eerola等人提出了一种基于期望的复杂性预测模型(EBM) ，他们发现音调 (由节奏位置和音高持续时间修正)、距离方向和距离变化的区间原则、节奏原则切片和节奏变化在预测听众的复杂性判断方面具有显著的能力。除了特征定义之外，模型构造还有算法选择。Cilibrasi等人质疑音乐聚类算法的有效性，并使用标准化的信息距离作为广义相似性度量来构建模型。在最近的一项研究中，Eerola等人使用信息理论和期望违背作为构建旋律复杂性的模型，并在七个数据集上验证了它，在这些数据集中，旋律复杂性的总体静态评分已经被收集和执行得很好。音乐评分的传统步骤是首先提取测试者的歌唱音乐，从歌唱输入信号中提取旋律特征，然后将提取的旋律特征与模板旋律进行相似性比较，根据它们之间的差异来衡量音乐表现。在旋律特征中，主要有三种类型的旋律特征: 参考频率、频谱系数的Mel和音高。针对不同的特征，不同的研究者提出了不同的提取方法，如利用快速傅里叶变换提取 MFCC 特征，利用 SHS 算法提取基音特征。

## 1.3 研究方法

本研究主要采用了文献分析法、案例分析法、问卷调查法、对比实验法。

**1.文献研究法：**文献分析法是指通过对某领域文献资料进行了收集、整理、筛选后，了解研究对象的性质和状况，并总结形成自己观点的分析方法。通过对视唱练耳教育心理学基础、视唱练耳游戏、动态难度调整进行资料查阅与分析，整理出视唱练耳游戏和动态难度调整二者结合的可行性；最后提出方案设计。

**2.案例分析法：**基于以上文献分析的基础与现有的游戏应用，可以找到相关学术或实践案例，通过对其进行研究与分析，总结其共性与差异、优势与劣势，为视唱练耳游戏的动态难度调整提供实例依据；

**3.问卷调查法：**指的是通过制定富有逻辑性、周密详尽的问题表格，来统一的收集广泛被研究群体的信息，从而对所研究的问题进行度量。本研究将在实验测试环节使用该方法，统一的对被试的情况进行信息收集，以方便后续的数据统计。

**4.对比试验法：**是指设置两个或两个以上的实验组，通过对比结果的比较分析，来探究各种因素与实验对象的关系。本研究将采用对照组的形式，在音乐学习者中进行实验，将接入了动态难度调整的游戏方案与非动态难度调整的游戏方案进行对比，从而探究其有效性和可行性。

## 1.4论文内容安排

### 1.4.1 研究思路

音乐是实施素质教育的重要内容和有效途径，对培养高素质人才有举足轻重的作用。据教育部统计，2019年，国内音乐类艺考报名人数约为14.3万人，2020年，这一数字达到16万人，增长率为11%。但同时，受利益驱使，音乐培训的师资质量良莠不齐，教育的规范性受到损害，专业音乐教育的高昂价格也使广大家长学生望而却步。视唱练耳作为琴童的必修课程，枯燥且具有较高的难度，往往需要花费大量的金钱或监护人的时间精力。但如没有此基础，将给未来的音乐学习埋下隐患。因此，本研究旨在增加学生对视唱练耳的学习动机、学习效果，提升音乐教育的规范性。

**1.解决视唱练习的普遍痛点。**视唱练习是指音乐专业的学生通过演唱乐谱上的音符(通常是一个几个音乐短句)来学习和提高音乐阅读能力的过程。刚开始学习音乐的学生经常被要求练习视唱，以建立他们对每个音符的音乐感知。视唱被认为是音乐表演和有效学习音乐知识的先决条件[4]。在视唱练习中，重要的是要有一个音乐专家的持续反馈，他能发现训练者犯的每一个错误，并找出纠正这些错误的最佳方法。传统上，视唱评估是在专家老师和学生之间一对一地进行的。然而，音乐教师的辨别力可能会受到主观因素和疲劳的影响。当学生在课外练习视唱时，很难从老师那里得到具体的指导和建议。因此，建立一个客观可靠的自动视唱训练应用是必要的。

**2.解决一些视唱练耳ICT方案的训练的评价缺失问题。**根据视唱学科知识的分类，视唱教学资源主要包括实例资源、视唱技能学习资源、音乐能力训练资源和教学录像资源。例子资源、视唱技能学习资源、音乐能力培训资源主要是各种类型的音乐文件，如五声图像、音乐模型音频、 MIDI 文件，用复杂的数据结构来描述和量化学习者的信息，以便计算机能够有效地识别学习者的学习特征。在智能教学系统中，学习者的个性特征是系统为他们提供个性化学习服务的主要决策基础。该算法通过分析学习者的特征数据来评估学习者的学习状态和能力水平，然后根据评估结果在既定规则下对学习资源和学习方法提供建议和指导学习者模型标准从一个总体上定义了学习者的特征

**3.开发一个针对视唱练耳特性的动态难度设计算法。**

本研究的重心在于研究视唱练耳训练的技能细分与动态难度设计算法。视唱练耳技能的学习相比商业游戏玩家技能的学习，有极大的差异。因此，在部分借鉴主流方法的研究基础上，需要开拓新的设计方法。

### 1.4.2研究对象

**1.探索一个新型的游戏化的视唱练耳训练方法。**游戏已被证明有对学习参与度、兴趣和成效的提升效果[5],相比传统的视唱练耳，游戏化的训练不仅节省了人力和工具，且游戏的高度交互性使得学习分析能够完美地捕捉学生的交互数据，从而更好地理解或改进学习过程。游戏的交互潜能可以增强学生的学习动机，从而提升训练效果。本研究旨在开发一个新型的视唱练耳训练游戏，并为学生提供便携、平滑、有趣的学习体验。

**2.** **给出有效的使用者学习状态评估**

根据系统的设计宗旨和特定学科学习者的特点进行细化、修改和定制，以便更好地为学习者提供个性化服务，提高学习效果。本文结合视觉歌唱学习者的特点，从旋律、音高、识谱、记忆方面设计视唱学习者模型。

**3. 选择已被验证的**OMLs**作为**DDA**方法的思想核心。**本研究建立在已有的方案基础上作进一步研究，结合游戏界被广泛采用的DDA(Dynamic difficulty adjustment)，对视唱游戏化的难度进行设计。对已有方案的难度体系进行调整，旨在提供更颗粒化的难度曲线，能够增加学生的学习动机和参与度，减少学习过程的阻滞感和枯燥感。以此游戏为载体的训练通过数据采集，借鉴自主调节

的OML(Open Models of Learning)模式，更好地反馈学生的弱点并提供建议[6]，提供更精细、更个性化[7]的训练。

### 1.4.3论文内容

# 第2章 视唱练习曲目难度评估研究

本章节主要通过对已有的常见训练曲目进行分析，构建曲目难度的测量方法，以进一步规划递进的训练路线。对于参与视唱练习的学生，音乐感知依赖于通过接触音乐而产生的内在心理模型。这种音乐依赖于两个认知过程: (1)统计学习，其中听众获得内部的认知模型的统计规律存在于他们所接触的音乐; (2)概率预测基于这些学习模型，使听众能够组织和处理他们的心理表征的音乐。为了证实这些假设，我回顾了一些研究，这些研究使用基于统计学习(音乐的信息动态模型)的概率预测计算模型来模拟人类听众的实证研究数据。结果表明，音乐感知涉及的广泛的心理过程——期望、情绪、记忆、相似性、分割和节拍——可以通过使用学习统计模型的一个单一的、潜在的概率预测过程来理解。此外，统计学习可以合理地预测不同文化暴露于音乐风格的因果效应，提供了一个文化距离的定量模型。理解音乐文化的神经基础将受益于经验神经影像学和潜在机制的计算模型之间的密切协调。

# 第3章 开放式学习模式研究

# 第4章 视唱练耳游戏设计研究与实现

## 4.1游戏流程设计

游戏流程采用OLM

## 4.2游戏可视化设计

基于多感官学习理论，游戏中对用户输入的视觉反馈，可显著提升沉浸式学习的效果。而多样的可视化方法可能产生学习效果上的差异。

### 4.2.1 音高匹配可视化

音高匹配，是认知心理学中的一类研究对象，即人类通过发声模仿指定的声调，这是视唱任务的重要组成部分，具有复杂的心理认知过程。根据已编码的音高序列准确发出音调依赖于听觉-声音系统的抽象逆模型，该模型允许歌唱者根据目标感知结果来计划喉部的运动姿势。研究表明，在这一过程中，听觉意向能够激活大脑的一些听觉和运动区域，以及次级听觉皮层和辅助运动区。这些区域的活动表明，大脑在执行不同的听觉意向任务，其中就包括操纵旋律的听觉图像。

根据MMIA模型，喉部运动和特定声音结果之间的关系模型，是构建于此个体过去的知觉事件。个人的一生中，统计学习使个体能够采集音高和喉部运动的关联经验，从而扩展形成抽象的模式。而多模态图像的示意图映射，允许个体将这些关联推广到新的情况。

因此，对于参与视唱练习的学生，及时的偏离反馈是必要的。这种反馈能够强化学生对于十二平均律内的音高认知。

音高在游戏内的具象化符号为玩家在天空中的位置。

## 4.2基音周期估计

AMDF(短时平均幅度差函数)，设语音信号的时间序列为,加窗分帧处理后得到的第帧语音信号为,其中下标表示第帧，设每帧帧长为。的短时平均幅度差函数定义为：

对于周期性的浊音语音，Di(k)呈现出了与浊音语音周期相一致的周期特性，利用这种特性，可以估计出浊音语音的基音周期。

在基于AMDF算法的基本提取中，将音频信号分为帧，得到的基音序列也是基于帧数刷新的。假设每位学习者的目标是准确地唱出乐谱，那么存在的野点不会持续很长时间，而且它们之前和之后应该有一个平滑的音高序列。事实上，通过分析大量视唱音频的音高序列中野点的持续时间，发现大量野点出现在两个平滑信号的中间，仅持续 1-2 帧。因此，我们可以 将音调序列平滑如下。

(1) 将基音序列中具有相等和相近基音值的帧作为一个子序列，并计算帧的数量。

(2)对每个基音序列的帧进行迭代，找到帧数在1到2之间且其前后帧大于2的野点。

(3)将发现的野点的音高值设置为其前后序列音高的均值。

(4) 将中值平滑应用于优化的基音序列。

## 4.3 视唱技能评估(Music Alignment)

DTW(Dynamic Time Warping)动态时间规整算法，是一种计算两个时间序列数据相似度的一种动态规划算法。

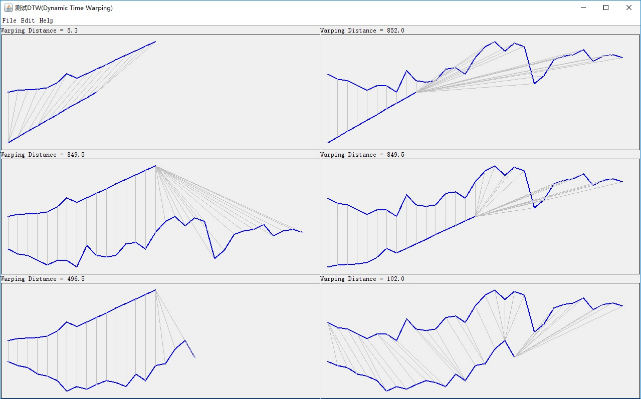
在实际任务中，用于计算相似性的两条数据在时间维度上是大致相似的，但很难得知每个时刻的对应关系。但此算法可以有效地解决整个难点。

图3.6 DTW算法示意图

上述DTW算法是针对所有时间序列的相似度计算而设计的，在本文的视唱评分中，可以结合数据的特殊性来改进基音序列的匹配与一般的语音信号匹配相比，视唱评分中基音序列的匹配过程具有以下两个特点。

(1) 视唱是指根据乐谱中的旋律唱出每个音符，乐谱中每个音符的持续时间远远长于基本检测中一帧的持续时间，通常呈乘法关系。这意味着从演唱音频中提取出来的音高序列由许多段组成，其音高值与序列相同。意味着当使用

DTW时, 基音序列相似度计算算法，可以通过子序列进行匹配，以指数形式减少算法的计算时间。 (2) 这里有三个方向用于选择原始 DTW 中的下一个路径点, 垂直选择表明序列 x 中的一个点可以对应于 y 中的多个连续点，反之亦然。在视唱评分中，假设视唱音频的音高序列为 x，模板的音高序列为 y，如果 y 中的多个连续点与 x 中的一个点对齐，则表示演唱者遗漏了一些音符，在这种情况下，y 中的这些点原来是错误的位置，在DTW 的路径选择中可以忽略，最后可以与相邻的点对应。换句话说，在最小累积距离路径选择中，可以跳过 y 中的点，这种改进也可以缩短计算时间。 一个人多次唱同一首歌，在个人音符和整体节奏上有所不同。因此，基音序列不能严格按时间匹配到模板，需要对基音序列进行时移、缩放等动态正则化操作。考虑到这一客观要求，本研究使用前面部分的DTW算法来实现模板匹配。

## 4.4 游戏流程

## 4.4 数据处理

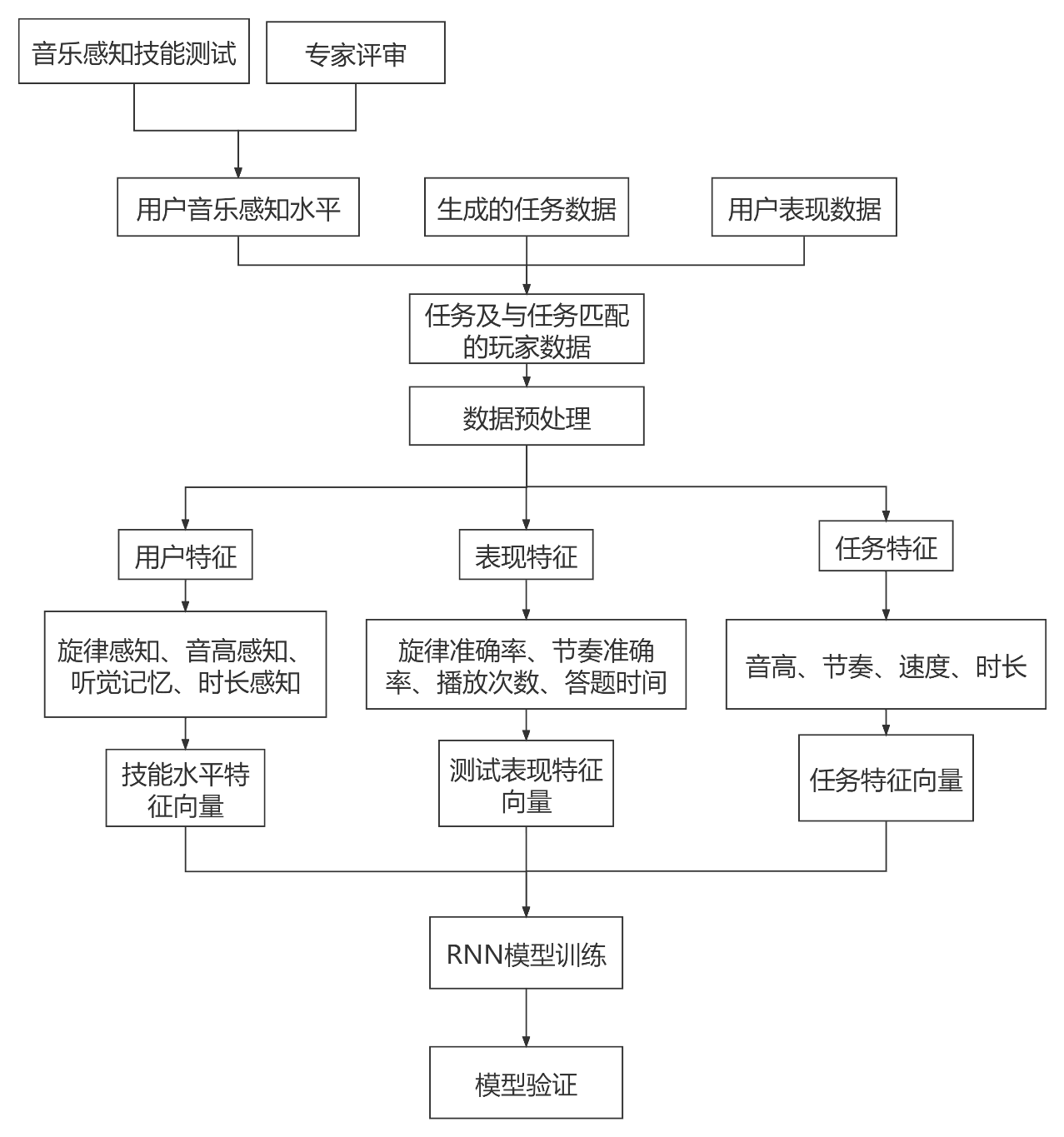


图3.7 数据处理流程

由于本游戏区别于大部分商业游戏的高频交互和基于时间的线性关卡体验，更偏向静态任务，我们采取

先生成任务与任务评分，再预测玩家表现的方式。在生成任务阶段，我们将随机生成音高任务和节奏任务。我们将利用算法生成大量的音高任务和节奏任务。通过标准测试和专家评审得出被试者的音乐感知水平，将被试者的感知水平和在任务中的表现，输入RNN深度神经网络进行训练，得出任务在多个维度上的难度评分。利用玩家个体在其中部分任务中的技能水平，结合任务的难度，预测其在任务中的测试表现，从而计算其在不同任务中的准确率预测值。

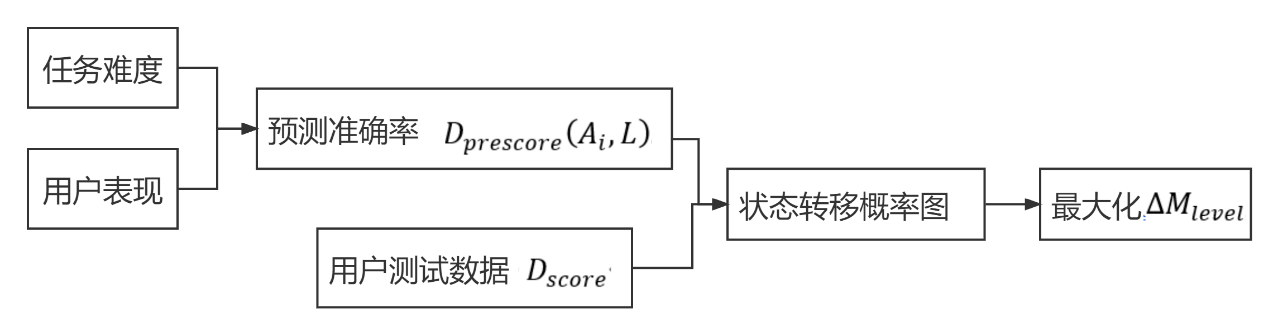
获取准确率预测值后，我们采用贝叶斯网络绘制玩家的状态转移概率图，通过动态规划方法求证最优的,以最大化玩家的水平提升值。依据MET音乐能力标准测试[49]结合玩家表现计算。

图3.8 求证最优水平提升值的数据流向图

3.9 自主调节难度的界面

上文提到，大部分的建模和适应过程都是难度调节系统在幕后进行的，学生们自己无法接触到。据进一步文献调查发现，跨不同类型系统的研究揭示了开放用户评估模型的好处，从而提高了系统透明度和用户参与度[50]。因此，在本研究中，我们希望尝试一种更加透明的适应性形式，除了提供游戏难度推荐之外，通过一个个性化的界面向学生显示相关的评估信息，并给予学习者自主调节的空间，如图3.9所示。

# 第5章 测试与数据分析

# 第6章 总结与展望

## 6.1本文工作总结

## 6.2研究创新点

1. 从**技术重组**的角度——提取DDA的核心思想即预测玩家表现与优化难度曲线，结合关卡自动生成与贝叶斯网络方法，以及视唱练耳任务本身的特点，强化了任务与提升视唱练耳提升用户技能水平目标的粘合度。

2.从**系统理论构建**的角度——本研究将DDA的核心思想迁移到教育游戏中，为音乐教育理论和游戏设计理论的结合形式做出探索；

3.从**教育媒介**的角度——本研究拓宽了音乐教育的形式，契合当下数字化时代的信息属性，改善了教育的媒介与途径。

4.从**游戏功能形式**的角度——本研究拓宽游戏在心理教育领域的功能性，让游戏充分发挥其轻量、即时、便捷、数字化的优势，把传统的、线下的教育改善至更方便快捷的形式。也为功能游戏在教育领域做出形式上的突破与创新实践。

参考文献

1. Habibi A, Damasio A, Ilari B, et al. Music training and child development: a review of recent findings from a longitudinal study[J]. Annals of the New York Academy of Sciences, 2018, 1423(1): 73-81.
2. Jaschke A C, Honing H, Scherder E J A. Longitudinal analysis of music education on executive functions in primary school children[J]. Frontiers in neuroscience, 2018, 12: 103.
3. North A C, Hargreaves D J, O'Neill S A. The importance of music to adolescents[J]. British journal of educational psychology, 2000, 70(2): 255-272.
4. D. Payne, “Essential skills for promoting a lifelong love of music and music making,” The American Music Teacher, vol. 54, no. 4, p. 26, 2005.
5. Syawaluddin A, Afriani Rachman S, Khaerunnisa. Developing Snake Ladder Game Learning Media to Increase Students’ Interest and Learning Outcomes on Social Studies in Elementary School[J]. Simulation & Gaming, 2020, 51(4): 432-442.
6. Romero C, Ventura S. Educational data mining and learning analytics: An updated survey[J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2020, 10(3): e1355.
7. Maseleno A, Sabani N, Huda M, et al. Demystifying learning analytics in personalised learning[J]. International Journal of Engineering & Technology, 2018, 7(3): 1124-1129.
8. N. J. Foss and T. Saebi, “Fifteen years of research on business model innovation,” Journal of Management, vol. 43, no. 1
9. S. Welch, R. Johns, G. Nicholls et al., “Incorporation of a theme-based research model into a pharmacy department research program,” Journal of Pharmacy Practice and Research, vol. 49, no. 3, pp. 262–269, 2019.
10. Tremblay G, Champagne F. Marking musical dictations using the edit distance algorithm[J]. Software: Practice and Experience, 2007, 37(2): 207-230.
11. Carrillo C, Baguley M, Vilar M. The influence of professional identity on teaching practice: Experiences of four music educators[J]. International Journal of Music Education, 2015, 33(4): 451-462.
12. Eyles A M. Teachers' perspectives about implementing ICT in music education[J]. Australian Journal of Teacher Education, 2018, 43(5): 110-131.
13. Debevc M, Weiss J, Šorgo A, et al. Solfeggio learning and the influence of a mobile application based on visual, auditory and tactile modalities[J]. British Journal of Educational Technology, 2020, 51(1): 177-193.
14. Pesek M, Vučko Ž, Šavli P, et al. Troubadour: A gamified e-learning platform for ear training[J]. IEEE Access, 2020, 8: 97090-97102.
15. Serdaroglu E. Ear Training Made Easy: Using IOS Based Applications to Assist Ear Training in Children[J]. European Journal of Medicine and Natural Sciences, 2018, 2(1): 61-68.
16. Hwang Y T, Chu C N. The Design of Music Ear Training System in Building Mental Model with Image Stimulus Fading Strategy[C]//International Conference on Learning and Collaboration Technologies. Springer, Cham, 2018: 127-135.
17. Fletcher C, Hulusic V, Amelidis P. Virtual reality ear training system: a study on spatialised audio in interval recognition[C]//2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games). IEEE, 2019: 1-4.
18. Wallentin M, Nielsen A H, Friis-Olivarius M, et al. The Musical Ear Test, a new reliable test for measuring musical competence[J]. Learning and Individual Differences, 2010, 20(3): 188-196.
19. Ryan R M, Deci E L. Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions[J]. Contemporary Educational Psychology, 2020, 61: 101860.
20. A. Stenberg and I. Cross, “White spaces, music notation and the facilitation of sight-reading,” Scientific Reports, vol. 9, no. 1, pp. 1–12, 2019.
21. E. Huovinen, “Early attraction in temporally controlled sight reading of music,” Journal of Eye Movement Research, vol. 11, no. 2, pp. 1–30, 2018.
22. J. C. Smith, “Hidden in plain sight: a music therapist and music educator in A public school district,” International Journal of Music Education, vol. 36, no. 2, pp.
23. H. Mitchell and B.. Roger, “'e moot audition: preparing music performers as expert listeners,” Research Studies in Music Education, vol. 39, no. 2, pp. 195–208, 2017.
24. Pathania A, Leiker A M, Euler M, et al. Challenge, motivation, and effort: Neural and behavioral correlates of self-control of difficulty during practice[J]. Biological psychology, 2019, 141: 52-63.
25. Tan C H, Tan K C, Tay A. Dynamic game difficulty scaling using adaptive behavior-based AI[J]. IEEE Transactions on Computational Intelligence and AI in Games, 2011, 3(4): 289-301.
26. Csikszentmihalyi M, Csikzentmihaly M. Flow: The psychology of optimal experience[M]. New York: Harper & Row, 1990.
27. Xue S, Wu M, Kolen J, et al. Dynamic difficulty adjustment for maximized engagement in digital games[C]//Proceedings of the 26th International Conference on World Wide Web Companion. 2017: 465-471.
28. Thai-Nghe N, Horváth T, Schmidt-Thieme L. Factorization models for forecasting student performance[C]//Educational Data Mining 2011. 2010.
29. Zook A E, Riedl M O. A temporal data-driven player model for dynamic difficulty adjustment[C]//Eighth Artificial Intelligence and Interactive Digital Entertainment Conference. 2012.
30. Spronck P, Sprinkhuizen-Kuyper I, Postma E. Difficulty scaling of game AI[C]//Proceedings of the 5th International Conference on Intelligent Games and Simulation (GAME-on 2004). 2004: 33-37.
31. Bull, S. (2020). There are open learner models about! IEEE Transactions on Learning Technologies., 13, 425– 448.
32. Dascalu, M.-I., Bodea, C.-N., Mihailescu, M. N., Tanase, E. A., & Ordoñez de Pablos, P. (2016). Educational recommender systems and their application in lifelong learning. Behaviour & Information Technology, 35(4), 290–297.
33. Bodily, R., & Verbert, K. (2017). Review of research on student-facing learning analytics dashboards and educational recommender systems. IEEE Transactions on Learning Technologies, 10(4), 405–418.
34. Aleven, V., McLaughlin, E. A., Glenn, R. A., & Koedinger, K. R. (2016). Instruction based on adaptive learning technologies. Handbook of Research on Learning and Instruction, 522–560.
35. Long, Y., & Aleven, V. (2016). Mastery-oriented shared student/system control over problem selection in a linear equation tutor. International Conference on intelligent tutoring systems, 90–100.
36. [14] L. Herrero and N. Carriedo, “'e role of cognitive flexibility and inhibition in complex dynamic tasks: the case of sight reading music,” in Current Psychology, pp. 1–13, Springer, Berlin, Germany, 2020.
37. [15]T. Herbert, “Public military music and the promotion of patriotism in the British provinces, c. 1780-c. 1850,” Nine-teenth-Century Music Review, vol. 17, no. 3, pp. 427–444, 2020.
38. [16] A. J. Bovin, “'e effects of frequent use of a web-based sightreading software on eighth graders’ music notational literacy,” Journal of Music, Technology and Education, vol. 11, no. 2, pp. 131–147, 2018. 131-147,2018.
39. [17] Ç. Ozen, “'e analysis and comparison of pre-service music teachers’ attitudes towards the piano lesson and their sightreading skills in terms of certain variables,” Journal of Education and Practice, vol. 8, no. 29, pp. 162–167, 2017.
40. [18] M. E. D¨uzbastılar, “'e comparasion of music teacher candidates’ instrument test grades and instrument sight reading grades in talent exam with individual instrument exam grades during four-years education period,” Journal of New Results in Science, vol. 17, no. 1, pp. 315–323, 2020.
41. [19] A. Adamyan, “Sight-reading as an important factor in the professional growth of future music teacher,” Revista Vortex, vol. 8, no. 2, pp. 1–16, 2020.
42. [20] Y. Liu, J. Wang, J. Li et al., “Zero-bias deep learning for accurate identification of Internet of things (IoT) devices,” IEEE Internet of 5ings Journal, vol. 84, p. 2627, 2634.
43. [21] S. Niu, Y. Liu, J. Wang et al., “A decade survey of transfer learning (2010–2020),” IEEE Transactions on Artificial In-telligence, vol. 1, no. 2, pp. 151–166, 2020.
44. [22] J. Yang, J. Wen, B. Jiang et al., “Blockchain-based sharing and tamper-proof framework of big data networking,” IEEE Network, vol. 34, no. 4, pp. 62–67, 2020.
45. [23] ] J. Yang, M. Xi, B. Jiang et al., “FADN: fully connected attitude detection network based on industrial video,” IEEE Trans-actions on Industrial Informatics, vol. 17, no. 3, pp. 2011–2020, 2021.
46. Stephenson M, Renz J. Agent-based adaptive level generation for dynamic difficulty adjustment in angry birds[J]. arXiv preprint arXiv:1902.02518, 2019.
47. Plass J L , Homer B D , Pawar S , et al. The effect of adaptive difficulty adjustment on the effectiveness of a game to develop executive function skills for learners of different ages[J]. Cognitive Development, 2019, 49:56-67.
48. Sampayo-Vargas S, Cope C J, He Z, et al. The effectiveness of adaptive difficulty adjustments on students' motivation and learning in an educational computer game[J]. Computers & Education, 2013, 69: 452-462.
49. Mikkel, Wallentin, and, et al. The Musical Ear Test, a new reliable test for measuring musical competence[J]. Learning & Individual Differences, 2010.

[50 ] Sarkar A, Cooper S. Meet your match rating: providing skill information and choice in player-versus-level matchmaking[C]//Proceedings of the 13th international Conference on the foundations of digital games. 2018: 1-8.