一种用于分类角色扮演游戏子类型的音乐信息检索方法

Daeun Hwang[1] **Xuyuan Cai**[1] **Edward F.Melcer**[1]**Elin Carstensdottir**[11] 计算媒体，加利福尼亚大学圣克鲁兹分校，美国dhwang8@ucsc.edu, xcai30@ucsc.edu

arXiv:2601.02591v1 [cs.SD] 5 Jan 2026

Late-Breaking / Demo Session Extended Abstract, ISMIR 2024 Conference

摘要

电子游戏音乐（VGM）通常采用与电影音乐相同的视角进行研究，研究重点在于其理论功能与媒体既定类型之间的关系。然而，迄今为止，我们尚未发现有系统的方法能够分析电子游戏音乐在多个既定游戏类型中的可量化音乐特征。因此，我们提取了三种角色扮演游戏（RPG）子类型中游戏的电子游戏音乐特征，并据此假设不同音乐特征与人们对各类型的认知和呈现方式之间存在相关性。这种相关性或可进一步表明，这些特征与该子类型所关联的预期叙事元素或游戏机制有关。

1. 引言

随着电子游戏媒体日益普及，电子游戏音乐（VGM）已在音乐产业中确立了独特的类型地位。电子游戏音乐通常与电影音乐一同进行研究，相关研究主要关注其理论功能与已识别的媒体类型之间的关系。尽管已有许多研究针对电影音乐采用系统的定量音乐信息检索方法 [1], 但针对电子游戏音乐与游戏类型之间关系的研究仍然不足。电子游戏音乐与传统电影配乐有很大不同，因为电子游戏与电影作为媒体形式存在根本差异 [2, 3]。

基于此，我们研究了涵盖三种已确定电子游戏子类型的配乐的音乐特征：冒险类角色扮演游戏、动作类角色扮演游戏和策略类角色扮演游戏。目前，游戏学者在进行游戏类型分类时，主要关注两个方面：游戏机制（游戏如何进行）和游戏叙事（游戏讲述的内容） [4]。事实上，任何媒介中的类型分类都与人类感知密切相关。尽管这两个方面对于游戏研究非常重要，

Hwang、X. Cai、E. Melcer 和 E. Carstensdottir。根据知识共享署名 4.0 国际许可协议（CC BY 4.0）授权。**署名：** D. Hwang、X. Cai、E. Melcer 和 E. Carstensdottir，《一种对角色扮演游戏子类型进行分类的音乐信息检索方法》，载于第 *25* 届国际音乐信息检索学会会议晚间演示环节扩展摘要，旧金山，美国，2024。

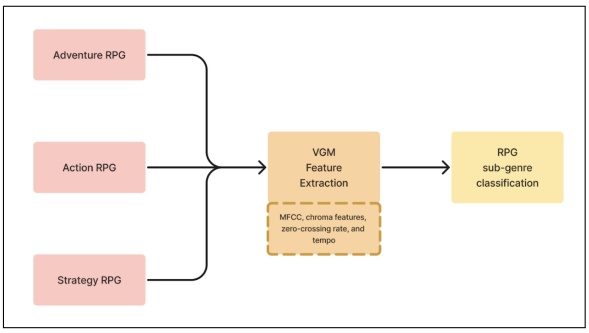

图1。角色扮演游戏子类型音乐特征表征概览

并且在进行游戏分析时能够提供可靠且易于理解的视角，但现有研究往往较少考虑声音和音乐。早期存储容量极小的游戏仍然加入了音效，甚至音乐 [3]。以如今的标准来看，游戏没有任何声音和音乐来伴随游戏玩法几乎是不可想象的。显然，音乐是电子游戏不可或缺的组成部分。通过对电子游戏音乐应用特征提取技术，我们可以研究电子游戏音乐的创作如何受到游戏类型识别的影响，并进一步促进游戏研究中对类型分类的研究。

2. 背景

随着时间推移，人们对音乐流派分类和特征提取的关注日益增加；特征提取是用于对流派进行分类和描述的关键方法。McKay 和 Fujinaga 指出了采用计算方法进行音乐类型分类的潜力 [5]。这种方法最初更多用于电影音乐，后来才应用于游戏音乐 [1]。采用特征提取进行音乐流派分类时，首先需要一个用于分析和训练的数据集 [6] 。用于描述音乐流派的特征类型多种多样，但最常用的是符号特征和音频特征 [7]。以往研究在符号特征和音频特征中采用了音高、音阶、梅尔频率倒谱系数（梅尔频率倒谱系数）等 [7–9]。然而，据我们所知，尚无研究聚焦于游戏研究中的音乐类型分类，这表明在使用计算方法研究游戏背景音乐方面仍然缺乏相

使用计算方法对游戏背景音乐进行检索。

3. 方法

考虑到可行性和易获取性，我们选择 Steam[1] 作为平台，根据受欢迎程度挑选这三个子类型的游戏。Steam 确定了角色扮演游戏中最受欢迎的三个子类型。选出每个子类型中最受欢迎的三款游戏后，我们从每款游戏中随机挑选三首器乐原声曲目，构建数据集。我们没有选择含歌词曲目，因为并非所有游戏中都能明显找到这类曲目，而纯音乐曲目是游戏原声音乐中更为主流的类型。我们总共收集了 27 首器乐原声曲目，分别来自 9 款不同的游戏；根据 Steam 的分类，这些游戏属于 3 种不同的角色扮演游戏子类型。

为了分析歌曲的特征，我们首先需要对数据集进行预处理。首先，将所有文件的音量归一化至 -5 dB，以统一分贝级别，然后剪辑为 15 秒。这 15 秒取自歌曲中间部分，以便大致捕捉歌曲的突出部分或主题。采样率设为 48000 Hz，以确保分析的一致性。最后，将所有文件转换为 WAV格式，以便在 Python 中进行特征提取。

主要使用 librosa 库 [10]的内置方法来计算以下变量：以每分钟节拍数（BPM）表示的平均速度、过零率（ZCR）、平均频谱质心、平均色度特征和梅尔频率倒谱系数。可通过链接[2] 访问 Google Colab 笔记本中的演示。

4. 结果

来自三个子类型歌曲的梅尔频率倒谱系数整体平均模式似乎没有显著差异。同时，冒险类角色扮演游戏的梅尔频率倒谱系数范围似乎相对最广，而策略类角色扮演游戏的范围最窄。

冒险类角色扮演游戏的平均过零率似乎比另外两个子类型的波动更大。这表明，冒险类角色扮演游戏的原声在音频信号方面可能更加丰富，这也暗示，与策略类角色扮演游戏和动作类角色扮演游戏相比，冒险类角色扮演游戏的游戏玩法具有相应的丰富性（例如，包含多个主题地点、多种情绪和多种游戏机制等）。

就平均频谱质心而言，角色扮演游戏的类型似乎并未造成显著差异。这表明，角色扮演游戏各子类型音乐的整体声音明亮度 [11]处于相近水平。

冒险类角色扮演游戏的平均色度特征似乎呈现出更为多样的旋律调性分布，与另外两个子类型相比，其中没有那么多突出的特定调性。这进一步印证了我们此前在过零率特征中观察到的模式：冒险类角色扮演游戏的旋律采用了多种调性，因此可能比策略角色扮演游戏和动作类角色扮演游戏更具适应性。

这三个子类型的平均速度相当接近，未观察到显著差异。一种可能的解释是，我们只使用了纯音乐曲目，而这类曲目的速度可能较为相似，与含歌词曲目相比尤其如此。

使用 K近邻模型[12], 进行训练和测试时，模型的准确率为 33.3%，不足以被认为是可靠的。这表明，可能需要使用其他算法进一步训练，或扩充数据集。

5. 局限性与未来工作

我们的研究仅关注角色扮演游戏及其子类型分类。虽然这为理解角色扮演游戏音乐的音乐特征表征提供了见解，但我们计划将研究扩展至更广泛的游戏类型。此外，在此数据集中纳入更多曲目，将有助于提高对不同游戏类型中电子游戏音乐的音乐特征进行分析时的准确性和具体性。

6. 结论

受电影音乐研究中音乐信息检索技术的启发，我们尝试以不同视角分析游戏音乐。通过提取音乐特征，我们将多种提取方法可视化，并发现了一些显著模式。尽管这未必能立即确立音乐特征与所识别类型之间的具体关系，但我们认为，观察到的音乐特征与所识别类型之间确实存在相关性，这本身就值得进一步研究。

7. 参考文献

[1] B. Ma、T. Greer、D. Knox 和 S. Narayanan，“从计算视角探究音乐如何刻画电影类型”，*PLOS ONE*，第16卷第4期，2021年。[2] K. Jørgensen，“该用新术语了吗？：重新审视电子游戏中的叙事内与非叙事内声音”，载于《游戏声音技术与玩家互动：概念与发展》。IGI Global，2011年，第78–97页。[3] K. Collins，《玩转声音：电子游戏中与声音和音乐互动的理论》。麻省理工学院出版社，2013年。[4] M. J. P. Wolf 编，《电子游戏媒介》。德克萨斯大学出版社，2001年。

[1]基于网络的视频游戏数字发行服务 https://store.steampowered.com [2]演示：https://colab.research.google.com/drive/1Fmwu6UKqbyrd-ih9WkgDxRu3ZDFBzvTm?usp=sharing

[5] C. McKay 和 I. Fujinaga，“音乐流派分类：值得研究吗？又该如何改进？”，载于*ISMIR*，2006年10月，第101–106页。[6] ——，“利用大型高层次音乐特征集进行自动流派分类”，载于*ISMIR*，第2004卷第2004期，2004年10月，第525–530页。[7] G. Tzanetakis、A. Ermolinskyi 和 P. Cook，“音频与符号音乐信息检索中的音高直方图”，《新音乐研究杂志》，第32卷第2期，第143–152页，2003年。[8] A. B. Mutiara、R. Refianti 和 N. R. A. Mukarromah，“使用支持向量机和音频特征进行音乐流派分类”，*TELKOMNIKA*（电信、计算与电子控制），第14卷第3期，第1024页，2016年。[9] M. Choudhury、R. Bhagwan 和 K. Bali，“宝莱坞音乐中旋律音阶的运用：一项实证研究”，载于国际音乐信息检索学会会议，2013年。[10]B. McFee、C. Raffel、D. Liang、D. P. W. Ellis、M. McVicar、E. Battenberg 和 O. Nieto，“librosa：使用 Python 进行音频与音乐信号分析”，载于第*14*届 *Python* 科学会议论文集，2015年，第18–25页。[11]J. M. Grey 和 J. W. Gordon，“频谱修改对音乐音色的感知影响”，《美国声学学会杂志》，第63卷，第1493–1500页，1978年。[12] G. Guo、H. Wang、D. Bell、Y. Bi 和 K. Greer，“基于K近邻模型的分类方法”，载于迈向有意义的互联网系统*2003*：*CoopIS*、*DOA*和*ODBASE*：*OTM*联合国际会议，*CoopIS*、*DOA*和*ODBASE 2003*，意大利西西里卡塔尼亚，*2003*年*11*月*3*日至*7*日。会议论文集。施普林格·柏林·海德堡，2003年，第986–996页。