通过计算方法理解人类对音乐抄袭的感知

DaeunHwang Hyeonbin Hwang加州大学圣克鲁兹分校　韩国科学技术院
dhwang8@ucsc.edu hbin0701@kaist.ac.kr

arXiv:2601.02586v1 [cs.SD] 5 Jan 2026

摘要

音乐相似性检测算法种类繁多，但现实世界中有关音乐抄袭的讨论往往基于听众感知。因此，我们旨在开展一项研究，考察人类感知音乐抄袭时所依据的关键标准，重点关注相似性分析中常用的三种音乐特征：旋律、节奏和和弦进行。在确定人们感知音乐相似性时使用的关键特征及变化程度后，我们提出一种“LLM 作为评审者”框架，采用系统化的逐步方法，并借助能够提取此类高层属性的模块。

1. 引言

音乐抄袭长期以来一直是个问题。这不仅可能损害音乐人的声誉，还可能对音乐的原创性与创造力提出质疑 [1, 2]。尽管人们已尝试多种方法来应对音乐相似性问题，但由于缺乏统一标准，音乐相似性检测算法的结果并不一致 [3, 4]。此外，现有的大多数音乐相似性研究方向略有不同：许多研究采用基于内容的音乐分析 [5–7], ，也有一些侧重于音乐推荐 [8, 9]。虽然音乐相似性检测有许多不同的算法，但只有人类专家才有可能全面评估这一复杂的多模态问题 [10]。媒体或社群中现有的多数关于音乐抄袭的讨论，都围绕着听众的感知展开。因此，本研究旨在从计算视角理解人们如何感知音乐相似性。我们还将基于影响人类音乐相似性感知的不同音乐元素权重，定义数据驱动的抄袭判定标准，并提出一个用于预测潜在抄袭案例的模型。

Hwang 和 H. Hwang。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 和 H. Hwang，《通过计算方法理解人类对音乐抄袭的感知》，载于第 *25* 届国际音乐信息检索学会会议晚期突破演示环节扩展摘要，美国旧金山，2024年。

2. 相关研究

现有文献显示，音乐相似性的定义与测量方式多种多样。许多研究采用多模态方法，通常涉及多个层级或空间的分析 [11, 12] ，或基于可视化的分析[2, 13]。现有研究通常会考察旋律 [1, 14, 15]、节奏 [16–19], 和和弦进行 [20–22]。旋律是音乐学中的核心概念，决定了歌曲的主旋律。旋律通常以音高为基础进行分析，或结合节奏加以分析。在本研究中，旋律模式将被视为由不同音高构成的音符序列，并通过半音表示的相对音高表示法呈现 [23], 而非绝对音高。音乐相似性语境下的节奏特征可定义为多个音符的时间序列和拍号 [19]。根据 Peeters 的说法，节奏特征主要基于五种方法之一进行分析：相似性矩阵、特征、时间模式、归一化周期性和声源分离 [24]。通常，这类分析会采用动态周期扭曲（DPW）度量来区分节奏相似性，因为该方法能减少遗漏节奏相似但速度不同的乐曲的情况 [17]。每个和弦都会给人以协和或不协和的印象 [25]，调性特征以及和弦的组合能够形成独特的情绪。因此，用于音乐相似性分析的和弦进行评估，将通过比较两个序列中每对和弦所计算出的和弦相似性相加得出 [22]。尽管音色广泛用于音乐相似性分析 [2, 11, 12], ，但本研究不会将其纳入考量，因为研究重点在于音乐的结构元素，而非整体声音特征或音质。由于音色指的是能够区分不同乐器的声音特质 [26, 27], ，为确保一致性，本研究将确保所有测试曲目使用相同的乐器音色。

3. 人类相似性感知研究

我们将首先开展一项研究，收集人类听众如何感知音乐相似性的数据。研究将以在线调查的形式进行，参与者将聆听每道题中嵌入的音频文件，并根据提示作答。我们将通过便利抽样招募参与者，包括具有音乐经验和非音乐经验的人士。年满18岁且听力能力没有障碍的参与者均有资格参加这项研究。

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

所有参与者在参加研究前都将获得知情同意书。我们将创建一个合成音乐数据集，其中包含不同程度的常用标准，这些标准来自以往关于音乐相似性和抄袭检测的研究：旋律、节奏和和弦进行。我们将使用生成的音乐曲目，而非现有歌曲，以便控制受测变量，从而避免混杂变量。此外，这也能减少版权问题，以及听众先前知识或偏好造成的偏差。生成的曲目将使用Logic Pro（一款 MIDI 软件）制作。在音乐领域专家的协助下，我们将首先生成5段曲目，每段约30秒，且和弦进行、旋律和节奏各不相同。所有曲目都将使用三角钢琴的相同默认音色录制。随后，我们将使用music21 [28], a（一款用于音乐脚本编写的 Python 库）处理这些原始曲目。为创建变体，我们将采用约0%、30%、50%和100%的变化幅度，生成不同的特征变化组合。排除完全未修改和完全修改的组合后，总计将得到62种变体。对于旋律变化，我们将使用移调函数，随机选取原曲中的音符进行改动。对于和弦进行，将采用多种技巧，例如和弦转位、和弦扩展和替代和弦。对于节奏变化，将调整音符时值和切分音。修改后曲目的速度也将进行调整。

参与者将会看到随机选取的15组音乐曲目对。他们需要判断自己是否认为这些曲目存在“抄袭”关系，以百分比对曲目之间的相似度进行评分，并简要说明作出判断的理由。研究人员将分析参与者的回答，以确定哪种特征及其变化程度产生了最显著的影响。此外，还将使用梅尔频率倒谱系数（MFCC）和色度特征等特征提取方法对曲目进行分析，以确定这些特征是否与参与者的感知之间存在显著相关性。

4. 预测模型

在界定音乐特征及相似程度的任何具有统计显著性的标准后，我们将训练一个模型，成功预测听众在抄袭情境下感知到的音乐相似性。正如先前研究所指出的，当前许多音乐相似性系统依赖低层特征，这可能导致有偏差的结果 [29]。鉴于此，我们提出一种基于以大语言模型为评判者的方法 [30]的新型框架，并赋予其专门的音乐知识。作为核心智能，大语言模型将利用其他模块提取的高级特征，例如和弦进行、旋律和节奏。随后，它将通过逐步的批判性思考来评估音乐相似性，这与人类听众运用关键要素和标准系统地评估音乐的方式相似。

5. 未来工作

为克服音乐相似性检测方法中的不一致性和多样性，本研究将验证听众对音乐抄袭的感知中存在的显著模式。在未来的工作中，我们还将全面梳理现有的音乐相似性检测算法，并酌情进行分类。通过分析调查问卷记录的听众标准，我们将考察听众判定抄袭时所考虑的具体标准，以及音乐特征变化的程度。我们将基于本研究分析得出的标准，采用以大语言模型为评判者的方法，开发用于相似性评估的预测模型。在后续研究中，可以考虑采用同时利用低层特征和高层特征的大语言模型评审模型，以实现更准确的预测。通过这项工作，我们旨在加深对人类音乐相似性感知的理解，并提供一种有助于检测抄袭风险的模型。

6. 参考文献

[1]Y. Li 和 X. Cheng，“关于文本指纹识别在旋律相似性检测中应用的研究”，载于*2024*年第五届人工智能、网络与信息技术国际研讨会（*AINIT*）。IEEE，2024年，第1109–1113页。[2] R. De Prisco、A. Esposito、N. Lettieri、D. Malandrino、D. Pirozzi、G. Zaccagnino 和 R. Zaccagnino，“音乐抄袭一览：相似性指标与可视化”，载于*2017*年第*21*届国际信息可视化会议（*IV*）。IEEE，2017年，第410–415页。[3] R. J. S. Cason 和 D. Müllensiefen，“唱同一首歌：计算旋律相似性测量与版权法”，《国际法律、计算机与技术评论》，第26卷，第1期，第25–36页，2012年。[4] A. Flexer，“音频音乐相似性中的评审者间一致性”，载于国际音乐信息检索学会会议（*ISMIR*）论文集，2014年，第245–250页。[5] B. McFee、L. Barrington 和 G. Lanckriet，“为音乐推荐学习内容相似性”，*IEEE*音频、语音与语言处理汇刊，第20卷，第8期，第2207–2218页，2012年。[6] J. T. Foote，“音乐与音频的基于内容的检索”，载于多媒体存储与归档系统*II*，第3229卷。SPIE，1997年，第138–147页。[7] P. Knees 和 M. Schedl，“基于音乐上下文数据的音乐相似性与推荐综述”，*ACM*多媒体计算、通信与应用汇刊（*TOMM*），第10卷，第1期，第1–21页，2013年。[8] B. Shao、D. Wang、T. Li 和 M. Ogihara，“基于声学特征和用户访问模式的音乐推荐”，*IEEE*音频、语音与语言处理汇刊，

与语言处理，第17卷，第8期，第1602–1611页，2009年。

[9] H. Han、X. Luo、T. Yang 和 Y. Shi，“基于特征相似性的音乐推荐”，载于 *2018*年*IEEE*安全生产信息化国际会议（*IICSPI*）论文集。IEEE，2018年，第650–654页。

[10] D. Müllensiefen 和 M. Pendzich，“关于音乐抄袭的法院判决及相似性算法的预测价值”，*MusicaeScientiae*，第13卷， 1_增刊，第257–295页，2009年。

[11] Y. Shiu、H. Jeong 和 C. C. J. Kuo，“利用相似性矩阵和动态规划进行音乐结构分析”，载于多媒体系统与应用 *VIII*，第6015卷。SPIE，2005年，第398–409页。

[12]D. Bogdanov、J. Serrà、N. Wack、P. Herrera 和 X. Serra，“统一低层次与高层次音乐相似性度量”，*IEEETransactions onMultime-dia*，第13卷，第4期，第687–701页，2011年。

[13] F. Simonetta、F. Carnovalini、N. Orio 和 A. Rodà，“基于图表示的符号音乐相似性”，载于*Audio Mostly 2018* 沉浸式声音与情感会议论文集，2018年，第 1–7 页。

[14] D. Müllensiefen 和 K. Frieler，“旋律相似性测量中的认知适切性：算法与人工判断”，音乐学计算，第 13 卷，第 2003年期，第 147–176 页，2004年。

[15]M. Casey 和 M. Slaney，“序列在音乐相似性中的重要性”，载于*2006*年 *IEEE* 国际声学、语音与信号处理会议论文集，第 5 卷。IEEE，2006年，第 V–V 页。

[16] D. P. Ellis、C. V. Cotton 和 M. I. Mandel，“用于音乐相似性的节拍同步表示互相关”，载于*2008*年 *IEEE* 国际声学、语音与信号处理会议。IEEE，2008年，第 57–60 页。

[17] A.Holzapfel 和 Y. Stylianou，《基于动态周期扭曲的音乐节奏相似性》，载于*2008*年 *IEEE* 国际声学、语音与信号处理会议。IEEE，2008年，第 2217–2220 页。

[18]T. Pohle、D. Schnitzer、M. Schedl、P. Knees 和 G. Widmer，《论节奏与总体音乐相似性》，载于国际音乐信息检索学会会议（*ISMIR*）论文集，2009年，第 525–530 页。

[19] E. Cao、M. Lotstein 和 P. N. Johnson-Laird，“音乐节奏的相似性与类别”，《音乐感知：跨学科期刊》，第31卷第5期，第444–469页，2012年。

[20] W. B. DeHaas、R.C.Veltkamp 和 F. Wiering，“调性音高步距：一种用于和弦进行的相似性度量”，载于国际音乐信息检索学会会议（*ISMIR*）论文集，2008年，第51–56页。[21] W. B. De Haas、M. Robine、P. Hanna、R. C. Veltkamp 和 F. Wiering，“音乐和弦序列相似性方法的比较”，载于《探索音乐内容：第七届国际研讨会，*CMMR 2010*，马拉加，西班牙，*2010*年*6*月*21*日至*24*日。修订论文 *7*》。施普林格·柏林·海德堡，2011年，第242–258页。

[22] C.Wongsaroj、N.Prompoon 和 A.Surarerks，《一种基于和弦进行与歌曲分段分析的音乐相似性度量》，载于*2014*年第四届数字信息与通信技术及其应用国际会议（*DICTAP*）。IEEE，2014年，第 158–163 页。

[23] J. S. Downie，《音乐信息检索》，信息科学与技术年度综述，第37卷第1期，第295–340页，2003年。

[24] G. Peeters，《用于音乐音频信号自动分类的节奏频谱与时间周期性表征》，《*IEEE*音频、语音与语言处理汇刊》，第19卷第5期，第1242–1252页，2010年。

[25] E. Aiba、K. Tobitani、T. Shimotomai、M. Tani、N. Nagata 和 T. X. Fujisawa，《基于和谐度的和弦特征评估模型：在音乐情绪可视化界面中的应用》，载于第六届软计算与智能系统国际会议暨第十三届先进智能系统国际研讨会。IEEE，2012年11月，第131–135页。

[26] R.Loughran、J. Walker、M.O’Neill 和 M.O’Farrell，《在乐器识别中使用梅尔频率倒谱系数》，载于*ICMC*，2008年8月。

[27] C. Fales，《音色悖论》，民族音乐学，第46卷，第1期，第56–95页，2002年。

[28] M. Cuthbert 和 C. Ariza，《Music21：用于计算机辅助音乐学和符号音乐数据的工具包》，载于国际音乐信息检索学会会议，2010年。

[29] A. Poltronieri，《基于知识的多模态音乐相似性》，*arXiv*预印本 *arXiv:2306.12249*，2023年。

[30] L. Zheng、W.-L. Chiang、Y. Sheng、S. Zhuang、Z. Wu、Y. Zhuang、Z. Lin、Z. Li、D. Li、E. P. Xing、H. Zhang、J. E. Gonzalez 和 I. Stoica，《使用 MT-Bench 和 Chatbot Arena 评判大语言模型评判者》，*arXiv*预印本 *arXiv:2306.05685*，2023年。