

# 弥合文化和数字鸿沟：面向全球南方公平音乐教育的低延迟 JackTrip 框架

Tiange Zhou

School of Future Design  
Beijing Normal University  
Zhuhai, China  
tiangezhoumusic@gmail.com

Marco Bidin

Department of Instrument Engineering  
Xinghai Conservatory of Music  
Guangzhou, China  
artdir.alea@yahoo.it

摘要—数字技术的迅速扩展已经改变了全球的教育格局，但在全球南部地区仍然存在显著的基础设施和文化挑战。本文介绍了一种低延迟的 JackTrip 框架，旨在弥合音乐教育中的文化和数字鸿沟。通过利用斯坦福大学 CCRMA 最初开发的一种基于开源、UDP 的音频流协议，该框架针对许多农村和地区技术限制进行了优化，如间歇性连接、带宽有限和高延迟等问题。研究系统地比较了 JackTrip 与 Zoom 等传统平台的表现，证明在模拟低资源条件下，JackTrip 能够实现低于 30 毫秒的延迟，并保持对非西方音乐传统至关重要的复杂音频细节。频谱分析证实，JackTrip 在处理微音阶、复杂数字节奏和和声纹理方面的卓越能力为实时合奏表演和音乐教育提供了一种文化上真实的媒介。这些发现强调了分散式边缘计算解决方案在赋予全球南部地区教师和音乐家权力方面所具有的变革潜力，促进了技术公平与文化保存。

*Index Terms*—数字鸿沟；JackTrip；低延迟音频流；网络音乐表演；音乐教育；全球南方；文化保存；边缘计算；去中心化架构；非西方音乐

## I. 介绍

音乐教育一直是文化传播的基石，特别是在那些以口头传统为主导的世界地区。在包括撒哈拉以南非洲、南亚和拉丁美洲在内的全球南方国家中，音乐远不止是一种艺术追求；它是身份认同、历史和精神的重要组成部分。例如，在西非，Griot 鼓手将历史事件刻录成复杂的多节奏模式，而在印度，古典音乐家则通过微分音程表演来表达深刻的情感和精神细微之处 [1]。这些音乐系统不能脱离其产生的文化背景，而表演通常会涉及音乐与非音乐的沟通，如口头讲故事和精神表现。

然而，这些丰富的音乐遗产正面临生存威胁，这在很大程度上是由现代技术和基础设施不足所引发的。根

据 Science Direct，在撒哈拉以南非洲 (SSA)，估计有 32% 的小学 and 近一半的中学仍然缺乏电力供应，无法实现互联网连接。[2]。这些基础设施缺陷严重阻碍了现代教育技术及数字平台的应用，并阻止它们促进全球艺术参与和文化保护的发展。

广泛可用的数字平台如 Zoom、Skype 和 Google Meet，虽然能够跨越地理距离连接世界各地的人们，但它们的设计更侧重于语音和一般沟通，而不是音乐表演的具体需求。这些平台基于优先考虑 12 平均律 (12-TET) 的音频标准设计，这种调音系统在西方音乐中很普遍，但对于许多全球南方音乐实践中使用的微分音、非西方音阶却不太合适。此外，这些平台存在高延迟和处理节奏复杂音乐的能力较差的问题，这使得它们不适合低资源环境中的实时合奏音乐教育。

针对如此严峻的挑战，本研究探讨了由斯坦福大学音乐与声学计算机研究中心 (CCRMA) 开发的低延迟 UDP 基音频流开源软件 JackTrip [3]。JackTrip 特别为实时网络音乐表演而设计，即使在高延迟和低带宽条件下也能正常工作，因此是促进全球南方地区远程音乐教育的理想选择。本研究从技术和文化两个角度评估了 JackTrip，探讨其支持非西方音乐表演、跨文化交流以及在全球南方资源匮乏地区的文化遗产保护与传播方面的潜力。

特别是，本研究指出了 JackTrip 在使全球南方地区的音乐合奏表演成为可能方面的潜力，在这些地区，互联网连接问题常常使得传统的在线音乐教育软件无法使用。通过利用廉价的边缘设备如 Raspberry Pi，

JackTrip 将处理任务从基于云的服务器中卸载出来，从而减少了延迟和成本。即使在农村和地区偏远的地方进行实时演出的能力，不仅对音乐教育具有重要意义，而且对于这些社区的文化交流和经济发展也具有重要影响。

## II. 相关工作

### A. 音乐教育在全球南方的重要性

全球南方的音乐教育不仅是艺术教学的一种方法，而且是维护文化认同和知识传承的关键组成部分。在世界上大多数地区，特别是撒哈拉以南非洲、南亚以及拉丁美洲的部分地区，音乐是社区生活的一部分，经常反映这些社区的历史、宗教和社会生活。例如，西非的格里奥传统深深植根于社会生活中，格里奥不仅扮演着音乐家的角色，还担任历史学家，通过音乐 [4] 编码和传递他们的人民的历史和知识。在这些社区中，音乐教育往往是一种口头传统，通过表演而非课堂教学从一代传给另一代。

这种教育音乐方法显然与西方教育体系不同，在西方教育体系中，正式的书面记谱系统和严谨的教学方法占主导地位。印度古典音乐教育基于复杂的拉格（旋律模式）和塔拉（节奏循环）系统，这些知识通过口述传承，并通过模仿和学徒制 [5] 学习。这些音乐传统不仅仅是艺术流派的代表，还是世界观、精神体系以及对人类状况理解的表现。这种文化特殊性强调了通过忠实于其本土形式和结构的教育来维护这些音乐传统的重要性。然而，全球南方的音乐教育面临着基础设施普遍不足所带来的严峻挑战。此外，在地理或经济距离限制了面对面教学的地方，假设面对面导师制和团体学习的传统教学模式也受到了压力。

然而，最紧迫的问题之一是在保持本土音乐传统与西方音乐体系日益普及之间进行的斗争。随着数字技术变得更加普遍，基于标准化记谱系统和音高集合（如 12-TET）的西方音乐教育模式在全球范围内占据了不均衡的优势地位。这种我们可称之为声波帝国主义的音樂霸权导致了非西方音乐实践被边缘化或扭曲。具体来说，Zoom 和 Skype 等在线学习的实际标准工具优化了语音和一般音频通信，但并未适应全球南方多样化的音乐形式。通过将调音系统标准化为西方 12-TET，这些平台压缩了非西方音乐传统中心的音调和节奏复杂性，例如印度古典音乐中的微音阶或非洲打击乐中错综复

杂的多节奏模式。这引发了一个重大问题：全球南方的音乐和学习体验被扭曲，在某些情况下甚至被抹去。

因此，这一挑战是双重的：如何在技术迅速发展的时代保护全球南方丰富多样的音乐遗产；以及如何发展能够尊重和培养这些文化遗产的音乐教育系统。日益扩大的数字鸿沟使这一问题变得更加紧迫，因为绝大多数来自全球南方的人被排除在外可以使他们融入全球音乐教育体系的技术进步之外。

### B. 在全球南方采用基于技术的音乐教育面临的挑战

全球南方在采用数字技术进行音乐教育方面面临的挑战不仅限于基础设施和文化保存问题。虽然人们非常关注硬件和互联网的访问问题，但同样重要的是影响新技术采纳的社会技术障碍。其中最显著的障碍之一是缺乏适合当地的技术、经济和文化条件的解决方案。例如，像 Zoom、Google Meet 和 Skype 这样的商业解决方案通常不适合音乐教育的具体需求。这些工具并非为实时音乐表演的复杂性而设计。它们优化了通用语音通信，并经常使用会降低音频质量的压缩算法，从而抹去了对音乐表达至关重要的细微差别。

伴随技术问题的还有数字素养的问题。在广大全球南方地区，优化使用在线音乐教育技术所需的技术知识存在显著差距。即使可以获取负担得起的技术，由于缺乏培训，这些技术往往被使用不足或误用。此外，许多在这些国家从事教学工作的人可能根本没有接受过与发达国家教师相同的教学培训或没有接触同等资源的机会，因此对于新的数字音乐教育技术的了解和实际应用都存在不足。

此外，存在一些经济问题使得广泛使用此类工具成为不可能。大多数低收入和中等收入国家的教育预算都很紧张，音乐和艺术教育总是最后才被分配资金。虽然智能手机和计算机等数字工具的成本已经降低，但缺乏稳定的电力供应、负担得起的互联网以及高昂的云服务成本都是其采用的重要障碍。

另一个关键考虑因素是能源获取的问题。全球南方农村和欠服务地区的电力中断频发，稳定供电的访问受限。国际能源署（IEA）估计撒哈拉以南非洲近 6 亿人没有任何电力供应，因此任何试图将基于技术的学习引入这些社区的努力都必须考虑到能源限制 [6]。在没有稳定电源的情况下，支持在线教育平台所需的设备如计算机、平板电脑和服务器都无法使用。因此，解决这一

困境的任何方案都必须以节能为设计重点,这使得低功耗设备如树莓派和太阳能供电项目特别有价值。

### III. JACKTRIP 作为资源受限环境下低延迟音乐表演的解决方案

与传统的基于 TCP/IP (传输控制协议/互联网协议) 和集中式服务器处理的云应用程序不同, JackTrip 的去中心化架构利用了边缘计算来最小化延迟并减少对高带宽互联网连接的依赖。这使其非常适合在低基础设施和高延迟环境中进行网络音乐表演,支持低于 50 毫秒延迟的实时音频传输,这是即使在网络高延迟环境下实现合奏音乐表演的基本要求 [7]。

与 Zoom 相比, JackTrip 允许在本地边缘设备上处理数据,而不需要中央服务器进行数据处理。这减少了对中央服务器的负担,并且可以使用低成本设备(如 Raspberry Pi) 来接收和发送音频信号。Raspberry Pi 设备价格低廉、能耗低且广泛可用,因此为资源有限的环境提供了一个很好的解决方案。通过利用这种低成本硬件, JackTrip 能够提供一种不依赖昂贵基础设施或云服务的廉价且可扩展的音乐教育解决方案。

JackTrip 的一个最重要的优点,特别是对于全球南方地区而言,是它可以保留非西方音乐传统的独特特征。JackTrip 开源的性质使其可以进行大量定制,包括将其软件调整为处理在南半球常见的替代调音系统、节奏和音乐结构。例如,印度古典音乐依赖于微分音阶和复杂的节奏周期,并需要一个能够处理音高和时间微妙之处平台,而像 Zoom 这样的平台并不具备这种处理能力。然而, JackTrip 允许用户通过使用 Pure Data (Pd) 等软件补丁来实现这些功能,这些补丁可以被调整为包含微分音阶或节奏敏感的缓冲算法 [8]。

除此之外, JackTrip 已经在一系列跨文化交流演出中证明了其文化相关性。例如,由 UCSD 组织的“Changing Tides”音乐会将韩国民间歌手和一位墨西哥笛手联合起来,利用 JackTrip 在地理位置上相距遥远的音乐家之间实现实时表演。这种合作展示了 JackTrip 不仅能促进合奏表演,还能作为文化交流的桥梁,让不同种类的音乐传统被全球观众听到并接触到。这种文化交[9]对于促进文化间的相互尊重和理解至关重要,尤其是在全球化倾向于导致少数西方音乐范式的主导地位 [9]。



图 1. “变幻的潮汐”电信音乐音乐会 2020 年 3 月。

尽管具有潜力, JackTrip 仍未在资源匮乏的环境中得到广泛应用。之前关于网络音乐表演的研究大多假设了带宽较高且延迟问题较少的环境(例如光纤网络)。本文试图通过评估 JackTrip 在支持全球南方地区实时音乐教育方面的潜力,并分析其技术与文化特性如何使其成为这些地区非西方中心音乐教育的理想工具,来填补这一空白。

### IV. 方法论

该研究采用了一种控制实验方法来评估和比较 JackTrip 和 Zoom 音频流平台在南半球农村地区典型网络条件下的有效性。模拟的条件包括有限带宽、高延迟和间歇性连接。两台计算机相距大约 1,000 公里,以代表远程教育机构之间的真实地理距离。使用 Network Link Conditioner 软件 (Mac OS) 和 Clumsy 软件 (Windows OS) 模拟了典型的 3G 和 4G 蜂窝连接的网络限制,并设置了带宽限制参数(上行: 5–10 Mbps, 下行: 10–50 Mbps) 和数据包丢失率 (0–3%) [10]。

音频通信的总延迟 ( $T_{total}$ ) 被数学地确定为三个主要组成部分的和:

$$T_{total} = T_d + T_p + T_{proc} \quad (1)$$

- **传输延迟 ( $T_d$ ):** 通过将数据包大小除以带宽来计算。假设音频流数据包大小为 512 字节 (4096 位) 和最小带宽 (5 Mbps), 传输延迟计算如下:

$$T_d = \frac{4096 \text{ bits}}{5 \times 10^6 \text{ bps}} \approx 0.819 \text{ ms}$$

- **传播延迟 ( $T_p$ ):** 通过将距离除以传播速度 (光纤中约为  $2 \times 10^8$  米/秒) 确定。对于 1,000 公里的距离:

$$T_p = \frac{1\,000\,000\text{ m}}{2 \times 10^8\text{ m/s}} = 5\text{ ms}$$

- **处理延迟 ( $T_{proc}$ ):** 通过计算机之间的多次 ping 测试实测得到, JackTrip 的延迟范围为 15-20 毫秒, 而 Zoom 的延迟范围为 130-140 毫秒。

来自两个平台的音频流在相同的网络条件下同时被记录。捕捉了三类音频样本:

- 1) 短时暂态声音 (小于 1 秒)
- 2) 中等长度的持续音符 (3-5 秒)
- 3) 延长持续时间的微分音阶用来自秘鲁的排箫演奏

使用 Audacity 软件进行了详细的频谱分析, 以评估幅度、频率和泛音内容。

## V. 结果

### A. 延迟分析

延迟计算突显了 Zoom 和 JackTrip 之间的显著性能差异:

- **总延迟时间:**

$$T_{\text{total}}(\text{Zoom}) = 0.819\text{ ms} + 5\text{ ms} + 135\text{ ms} \approx 141\text{ ms}$$

- **杰克特里普总延迟:**

$$T_{\text{total}}(\text{JackTrip}) = 0.819\text{ ms} + 5\text{ ms} + 20\text{ ms} \approx 26\text{ ms}$$

JackTrip 表现出显著更低的延迟 (大约 26 毫秒), 这适用于实时合奏表演。相比之下, Zoom 较高的延迟 (大约 141 毫秒) 对实时音乐同步产生了负面影响。

### B. 音频信息分析

频谱分析清楚地识别出 Zoom 和 JackTrip 之间存在显著差异:

- **缩放:** 显著的音频压缩降低了清晰度, 特别是在 3 kHz 以上, 负面影响了音乐细节的准确性和保真度。
- **杰克特里普:** 保留全频段音频细节 (20 Hz 到 20 kHz), 保持微分音准确性、节奏复杂性和和声丰富性。

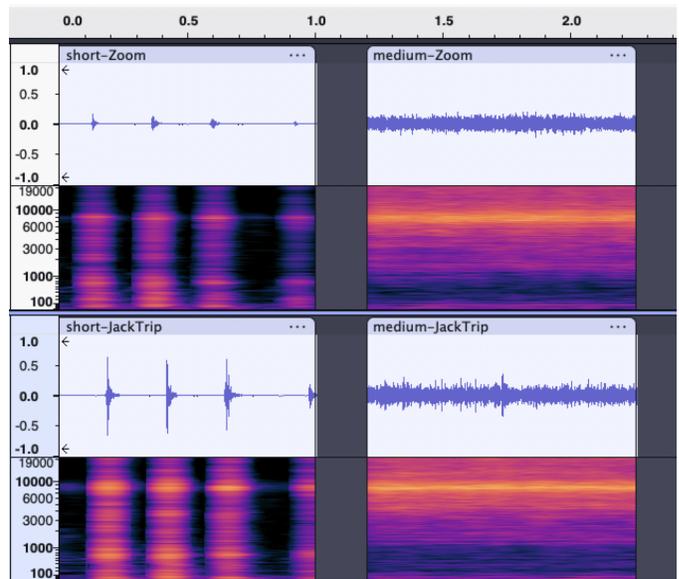


图 2. 短时间和中时长声音通过 Zoom 和 JackTrip 的比较分析

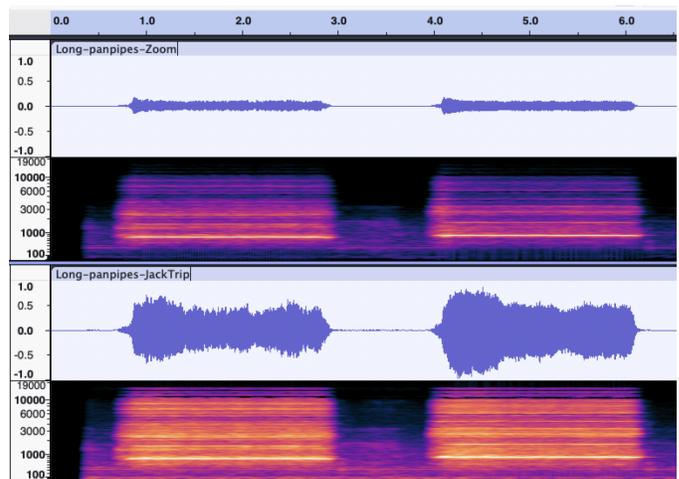


图 3. 长笛长时间声音通过 Zoom 和 JackTrip 的比较分析

### C. 音频质量和文化忠实度

JackTrip 录音有效地保留了微分音准确性、节奏复杂性和和声真实性, 这些都是准确表现非西方音乐传统所必需的。相比之下, Zoom 的压缩音频减少了这些重要的音乐细节, 从而削弱了文化真实性。这些结果强调了 JackTrip 在技术支持有限的全球南方背景下支持多样且细腻的音乐传统的技术和文化相关性。

## VI. 讨论

### A. 技术表现及其在全球南方音乐教育中的适用性

本研究展示了 JackTrip 与 Zoom 在延迟、数据包丢失以及节奏和音调准确性方面的关键音频信息差异

——这些因素在网络条件类似于远程南半球地区所遇到的情况下尤为重要。JackTrip 提供的低延迟实时合奏表演为南半球音乐教育中的电子学习提供了一个可能具有变革性的解决方案。Zoom 中平均音频延迟约为 115–141 毫秒，超出了实时人类音乐连接的阈值，而 JackTrip 平均约 25 毫秒的延迟则允许音乐家和教育者即使在地理位置遥远的情况下也能进行实时协作。此外，通过 Audacity 中的详细频谱分析证实，JackTrip 提供 1536 千比特每秒的音频质量——相比之下 Zoom 为 192 千比特每秒。

### B. 文化相关性：保持非西方音乐传统

JackTrip 的文化影响超越了其技术能力。不充分的音频传输可能导致简化解释和文化完整性的侵蚀。通过实现高质量、低延迟的音频通信，JackTrip 使全球南方的音乐家和教育者能够保持本地音乐实践，包括复杂的节奏、微分音细微差别和即兴传统，而不是让他们的传统被以西方为中心的音乐理论（如 12-TET）过度简化。

### C. 赋能全球南方：教育与社会转型

虚拟音乐教育作为全球南方的一项创新因素，其力量在于对社会、经济和文化的根本性变革潜力。在传统基于语音的音乐教育由于基础设施和经济限制而罕见的地区，JackTrip 为学生和教师之间的合作提供了一个新的前沿领域，无论他们的地理位置或财务状况如何。通过培养全球教育网络，JackTrip 可以作为促进社会变革的催化剂，推广文化多元主义和音乐教育的民主化。

### D. 挑战与未来方向

虽然这项研究为 JackTrip 在提升全球南方音乐教育方面的潜力提供了坚实的基础，但仍存在挑战。低成本、节能硬件如树莓派的实施受到诸如电力供应、网络连接和维护技能等限制。此外，整合数字技术必须以补充而非消除传统社区学习实践的方式进行。未来的研究应包括在当地社区进行试点测试、共同开发具有文化相关性的教学方法以及对数字技术对传统音乐教育实践影响的纵向研究。

## VII. 结论

本研究证实，JackTrip 这款开源低延迟音频流软件是全球南方网络音乐教育的有效工具。通过在模拟网络条件下对 JackTrip 和 Zoom 进行比较评估，研究

揭示 JackTrip 提供显著更低的延迟、更好的音质以及增强的文化真实性。因此，即使是在资源匮乏的环境中，JackTrip 也能促进实时合奏表演，而传统的平台如 Zoom 却无法提供必要的功能。

JackTrip 高容错率、低延迟以及对复杂节奏和微分音调的支持使其非常适合非西方音乐传统的精细要求。其开源灵活的架构允许全球南方的地方教育工作者和音乐家根据自己的特定文化需求调整平台。最终，JackTrip 不仅代表着技术上的创新，也代表着文化的创新，通过公平获取全球音乐教育的机会来弥合数字鸿沟并赋予边缘化社区力量。

## 参考文献

- [1] Y. Winter, "On the grammar of a Senegalese drum language," *Language*, vol. 90, no. 3, pp. 644 – 668, Sep. 2014.
- [2] M. Moner-Girona, F. Fahl, G. Kakoulaki, D.-H. Kim, I. Maduako, S. Szabo, G. Nhamo, B. K. Sovacool, and D. J. Weiss, "Empowering quality education through sustainable and equitable electricity access in African schools," *Joule*, vol. 9, no. 2, pp. 345–373, Feb. 2025.
- [3] M. Bosi, A. Servetti, C. Chafe, and C. Rottondi, "Experiencing remote classical music performance over long distance: a JackTrip concert between two continents during the pandemic," *Journal of the Audio Engineering Society*, vol. 69, no. 12, pp. 934 – 945, Dec. 2021.
- [4] S. A. Ebine, "The Roles of Griots in African Oral Tradition among the Manding," *\*International Journal of Research Development\**, vol. 11, no. 1, pp. 1 – 10, 2019.
- [5] R. K. Wolf, Ed., *Theorizing the Local: Music, Practice, and Experience in South Asia and Beyond*. Oxford, U.K.: Oxford University Press, 2009.
- [6] International Energy Agency, "Electricity access continues to improve in 2024 - after first global setback in decades," IEA, Paris, 2024. [Online]. Available: <https://www.iea.org/commentaries/electricity-access-continues-to-improve-in-2024-after-first-global-setback-in-decades>.
- [7] P. Cairns, H. Daffern, and G. Kearney, "Investigation of Server – Based Spatial Audio for Metaverse Concert Distribution," in *Proc. 2024 IEEE 5th Int. Symp. Internet of Sounds (IS2)*, Erlangen, Germany, Sep. 30 – Oct. 2, 2024, pp. 1 – 8.
- [8] A. T. Porres, "Dissonance Model Toolbox in Pure Data," in *Proc. 4th Pure Data Convention*, Weimar, Germany, 2011.
- [9] M. Dessen, "all nearness pauses, while a star can grow," eScholarship, 2020. [Online]. Available: <https://escholarship.org/uc/item/50j6x5cr>.
- [10] J. Du, C. Zhang, T. Tang, and W. Qu, "Learning-based transport control adapted to non-stationarity for real-time communication," *2024 IEEE/ACM 32nd International Symposium on Quality of Service (IWQoS)*, pp. 1 – 10, 2024.