



OPEN ACCESS

EDITED AND REVIEWED BY
Alessandra Sciutti,
Italian Institute of Technology (IIT), Italy

*CORRESPONDENCE
Huijiang Wang,
✉ hw567@cam.ac.uk

RECEIVED 07 February 2025
ACCEPTED 14 February 2025
PUBLISHED 25 February 2025

CITATION
Wang H, Hughes J, Nonaka T, Abdulali A,
Lalitharatne TD and Iida F (2025) Editorial:
AI-powered musical and entertainment
robotics.
Front. Robot. AI 12:1572828.
doi: 10.3389/frobt.2025.1572828

COPYRIGHT
© 2025 Wang, Hughes, Nonaka, Abdulali,
Lalitharatne and Iida. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC
BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Editorial: AI-powered musical and entertainment robotics

Huijiang Wang^{1*}, Josie Hughes², Tetsushi Nonaka³,
Arsen Abdulali¹, Thilina Dulantha Lalitharatne⁴ and Fumiya Iida¹

¹Bio-Inspired Robotics Lab, Department of Engineering, University of Cambridge, Cambridge, United Kingdom, ²CREATE-Lab, Department of Mechanical Engineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, ³Graduate School of Human Development and Environment, Kobe University, Kobe, Japan, ⁴School of Engineering and Materials Science, Queen Mary University of London, London, United Kingdom

KEYWORDS

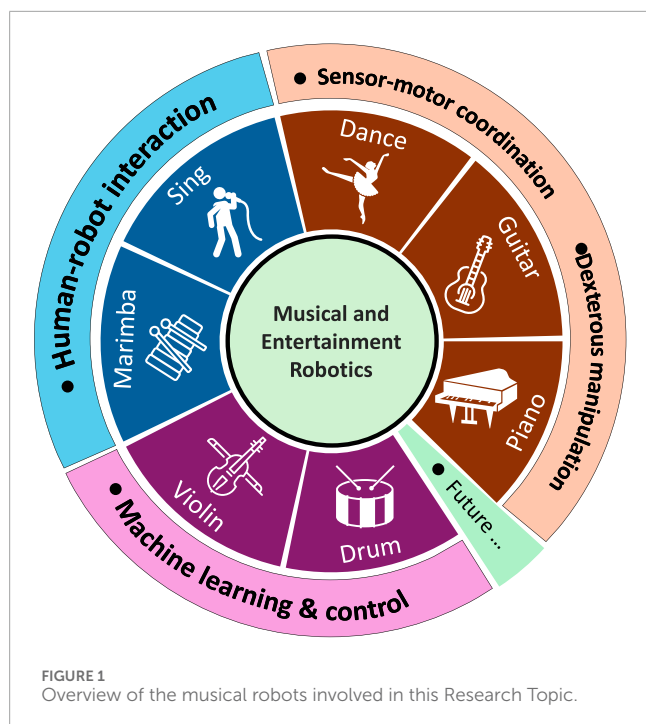
human-robot interaction, dexterous manipulation, musical and entertainment robots, machine learning, wearable devices, robotic expressiveness

Editorial on the Research Topic AI-powered musical and entertainment robotics

The convergence of robotics and artificial intelligence (AI) is revolutionizing the field of music and entertainment. Robots are evolving from performing traditional service-oriented tasks to enabling advanced human-robot interaction (HRI) with potential emotional engagement. The pursuit of robotic expressiveness presents new challenges and opportunities in the modeling, design and control of musical and entertainment robots. Current studies mainly work on the design and physical implementation of robots capable of manipulating various musical instruments (Wang et al., 2022; Lim et al., 2012), while the development of socially intelligent robots for real-time HRI remains underexplored. With advancements in AI, robots can now compose and improvise, as well as interpret and respond to human affective states during HRI (McCull et al., 2016; Wang et al., 2024).

This Research Topic was initiated to present the latest developments of AI-powered musical and entertainment robots. As a result of the call, six papers have been accepted and collected in this Research Topic. These articles provide a comprehensive exploration of diverse artistic forms including singing, dancing and musical performance on instruments such as the piano, violin, guitar, drum and marimba. Figure 1 shows an overview of the musical robots investigated in these studies.

Among the contributed works, two articles focused on dexterous manipulation and sensorimotor coordination. Gilday et al. introduced a general-purpose system featuring a parametric hand capable of playing both the piano and performing guitar pick strumming. Unlike existing bespoke robotic musical systems, the proposed hand was designed as a single-piece 3D-printed structure, demonstrating potential for enhanced expressiveness in entertainment applications through the modulation of mechanical properties and actuation modes. The study highlighted that leveraging system-environment interactions enabled diverse, multi-instrument functionalities and variable playing styles with simplified control. Instead of musical instrument playing, Twomey et al. investigated dance performance using wearable soft sensors on the arm to explore whether such devices could enhance artistic expression. Dance movements were modeled as colliders within virtual mass-spring-damper systems, and limb segments were analyzed in local frames to avoid drift issues commonly associated with IMUs. The authors proposed a parallel algorithm to detect



improvisational dance movements and control soft wearable actuators which can change size and lighting in response to detected motions. This work exemplified sensorimotor coordination and demonstrated how traditional dance and aesthetics could be enriched by spontaneous wearable-driven movements.

Robot learning and control represent one of the biggest challenges in musical and entertainment robotics, particularly for acquiring manipulation skills and robotic expressiveness. Horigome and Shibuya developed a RL-based controller for a violin-playing robot, a 7-DoF dual-arm system actuated by DC motors. The system mimics human performance with the left arm handling fingering and the right arm controlling bowing movements. The right arm regulates multiple parameters including bowing speed, pressure, sounding point and direction. Analysis of the target sound pressure demonstrated that the robot successfully learned violin-playing techniques and enables expressive performance variations. The robot was automated to play the violin based on musical scores, demonstrating its ability to interpret and execute complex musical tasks. Similarly, Karbasi et al. explored robotic drumming using a two-DoF robotic arm with flexible grippers, which is referred to as ZRob. They employed an RL-based algorithm with a Deep Deterministic Policy Gradient (DDPG) architecture, incorporating both extrinsic and intrinsic reward signals. The results showed that intrinsic rewards triggered the emergence of novel rhythmic patterns. Additionally, the robot's physical dynamics—embodied intelligence—were found to influence the learning algorithm due to the physical constraints of the drumming setup. This study highlights the interplay between robotic hardware and learning algorithms in achieving expressive musical performance. It can be seen that reinforcement learning continues to be a powerful and widely utilized approach for enabling robots to acquire complex manipulation and expressive skills.

The aforementioned studies have investigated both hardware and software advancements. However, the interaction between these robots and humans has not been explored. Gao et al. investigated synchronization between human musicians and Shimon, a robotic marimba player capable of head and arm movements. Their study revealed that ancillary and social gestures, particularly head movements, significantly enhance temporal synchronization between humans and robots. Through experiments with human participants, the results demonstrated positive social engagement when collaborating with robots in artistic performances. The study also found that social head gestures improved synchronicity slightly more than ancillary or instrumental gestures, providing quantitative insights into the role of non-verbal cues in HRI. Similarly, Nishiyama and Nonaka investigated the concept of “togetherness” in a singing scenario, where human participants coordinated their voices with either another human or a machine (Vocaloid) under non-visual conditions. The study highlighted that human-to-human cooperation achieved higher similarity and anticipatory synchronization compared to human-machine interaction. These findings highlight the critical role of embodiment in enabling natural and effective collaboration, demonstrating how physical presence and human-like traits shape interaction dynamics.

In conclusion, reinforcement learning holds strong potential in tackling the key challenges of equipping musical robots with advanced skills. Current AI-driven robotic systems have demonstrated the feasibility of achieving robotic expressiveness in various musical instruments. However, human-robot interaction presents a more complex Research Topic that requires interdisciplinary collaboration across fields such as robotics, materials science, computer science, psychology, musicology, sociology and ethics.

Author contributions

HW: Project administration, Visualization, Writing—original draft, Writing—review and editing. JH: Writing—review and editing. TN: Writing—review and editing. AA: Writing—review and editing. TL: Writing—review and editing. FI: Supervision, Writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the SMART project, an EU Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860108.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

References

Lim, A., Ogata, T., and Okuno, H. G. (2012). Towards expressive musical robots: a cross-modal framework for emotional gesture, voice and music. *EURASIP J. Audio, Speech, Music Process.* 2012, 3–12. doi:10.1186/1687-4722-2012-3

McColl, D., Hong, A., Hatakeyama, N., Nejat, G., and Benhabib, B. (2016). A survey of autonomous human affect detection methods for social robots engaged in natural hri. *J. Intelligent & Robotic Syst.* 82, 101–133. doi:10.1007/s10846-015-0259-2

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Wang, H., Howison, T., Hughes, J., Abdulali, A., and Iida, F. (2022). "Data-driven simulation framework for expressive piano playing by anthropomorphic hand with variable passive properties," in *2022 IEEE 5th international conference on soft robotics (RoboSoft)* (IEEE), 300–305.

Wang, H., Zhang, X., and Iida, F. (2024). Human-robot cooperative piano playing with learning-based real-time music accompaniment. *IEEE Trans. Robotics* 40, 4650–4669. doi:10.1109/tro.2024.3484633