Learning to Move with Style: Few-Shot Cross-Modal Style Transfer for Creative Robot Motion Generation

Kieran Woodward
School of Computer Science
University of Nottingham
Nottingham, UK
kieran.woodward1@nottingham.ac.uk

Alicia Falcon-Caro
School of Medicine
University of Nottingham
Nottingham, UK
alicia.falconcaro@nottingham.ac.uk

Steve Benford
School of Computer Science
University of Nottingham
Nottingham, UK
steve.benford@nottingham.ac.uk

Abstract

As robots increasingly participate in creative and social contexts, the ability to generate creative, stylised movements becomes crucial for applications ranging from performance art to human-robot collaboration. We present a novel framework for cross-modal style transfer that enables robots to learn new movement styles by adapting existing human-robot dance collaborations using human movement videos. Our dual-stream architecture processes raw video frames and pose sequences through cross-modal attention mechanisms, capturing rhythm, acceleration patterns, and spatial coordination characteristics of different movement styles. The transformer-based style transfer network generates motion transformations through residual learning while preserving the trajectory of original dance movements, enabling few-shot adaptation using only 3-6 demonstration videos. We evaluate across ballet, jazz, flamenco, contemporary dance and martial arts, introducing a creativity parameter that provides control over the style-trajectory trade-off. Results demonstrate successful style differentiation with overall style transfer scores increasing 6.7x to 7.4x from minimum to maximum creativity settings, advancing human-robot creative collaboration by expanding robots' expressive vocabulary beyond their original choreographic context.

1 Introduction

In recent years, the convergence of Artificial Intelligence (AI), robotics and creativity has created new opportunities for human-robot interaction in creative contexts (1). The Different Bodies project explored how professional dancers could engage with robots through direct physical touch. It undertook a series of dance improvisations in which one dancer physically manipulated a Franka Panda robot arm while a second responded to the resulting movements as they were mirrored live on a second arm. This yielded insights into how humans can safely, ethically and creatively interact with robots at close quarters (2; 3; 4). It also raised the idea that, based on the examples the dancers had created, the robots might learn to generate their own suitably stylistic movements in order to become more active co-performers as we explore in this paper.

Ethical concerns around AI in creative fields has raised questions about the replacement of human artists (5; 6; 7). This has encouraged research towards more ethical AI generation of artworks,

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Creative AI Track.

where artists remain part of the creative process (8; 9; 10; 11). Interactive creative generation, or co-creative artwork generation, has emerged as a promising approach, showcasing possibilities where artistic expertise directs generative models whilst maintaining the artist's style (12; 1). Our approach embodies this co-creative paradigm by enabling robots to stylise movements from human dancers.

On the other hand, generative models have expanded beyond the visual arts to explore music generation (13), painting (14; 15), and dance generation (16; 17). Furthermore, advances in interactive adaptation of generative models have demonstrated the potential for real-time style transfer in creative contexts (18; 19). Meanwhile in the last decade, robots are increasingly moving beyond purely functional tasks towards domains where expressivity, aesthetics, and social interaction play a central role (20; 21). This evolution is particularly relevant for dancers who may find robotic collaboration for artistic expression that complement and extend their embodied capabilities.

In the artistic robotic field, a key component of this integration lies in enabling robots to produce stylised movements that are not only physically feasible, but also reflect particular artistic or cultural styles such as tango, jazz or ballet (22). Such capabilities can deepen human-robot interaction, contribute to digital performance art and support embodied AI systems that operate in creative or social settings. Furthermore, robots capable of performing stylised dance or artistic movements can act as co-creators or co-dancers in performances, improving and enhancing the artistic expression.

While prior work in robotics has explored movement stylisation in contexts such as character animation or motion retargeting (23; 24), most current systems rely on pre-programmed sequences, or lack the ability to adapt stylistic features to new motion contexts dynamically, especially in real-time or with limited training data. Therefore, fast robot motion adaptation to new styles remains a major challenge.

Recent advances in machine learning, particularly deep generative models, have provided tools for learning style representations from motion data (25; 26). In robotics, stylised motion synthesis has been approached through learning from demonstration (27; 28), and reinforcement learning (29; 30). However, these approaches often target high-level attributes rather than fine-grained artistic styles and typically require extensive training data that may not be available in creative contexts.

In this paper we propose a novel framework for adapting an existing robot motion to human movement styles through cross-modal style transfer. Our proposed approach enables robots to learn expressive movement from few video demonstrations, requiring only 3-6 examples to adapt to new styles such as ballet, jazz, or martial arts. The system processes both raw video frames and pose keypoint sequences through a dual-stream architecture with cross-modal attention, capturing temporal dynamics, rhythm, and spatial coordination that characterise different movement styles. By leveraging few-shot adaptation, our method is practical for real-world creative applications creating new possibilities for artistic expression and human-robot co-creation.

The remainder of this paper is organised as follows, Section 2 explains the proposed technique, Section 3 discuss the results obtained along with the limitations of the proposed method and Section 4 concludes. The code which describes all parameters that contribute to the performance of the proposed method, is available at https://github.com/kieransfwoodward/Learning-to-Move-with-Style-Few-Shot-Cross-Modal-Style-Transfer-for-Creative-Robot-Motion-Generation

2 Motion Style Transfer System Architecture

The proposed robot motion style transfer system employs a dual-stream architecture that combines visual motion analysis with pose-based movement understanding to transfer human movements to existing robotic movements while preserving their original trajectory characteristics. The system comprises of a dual-stream temporal encoder that processes raw video frames and extracted pose data, a robot motion encoder which captures the motions of the original robotic trajectories, a transformer-based style transfer network that generates motion transformations through residual learning and an adversarial discriminator ensuring motion realism. Overall, this novel architecture utilises multi-modal fusion through cross-modal attention mechanisms, temporal transformer blocks and adversarial training to achieve few shot adaptation to new movement styles using only 3-6 demonstration videos. Fig. 1 summarises the proposed method.

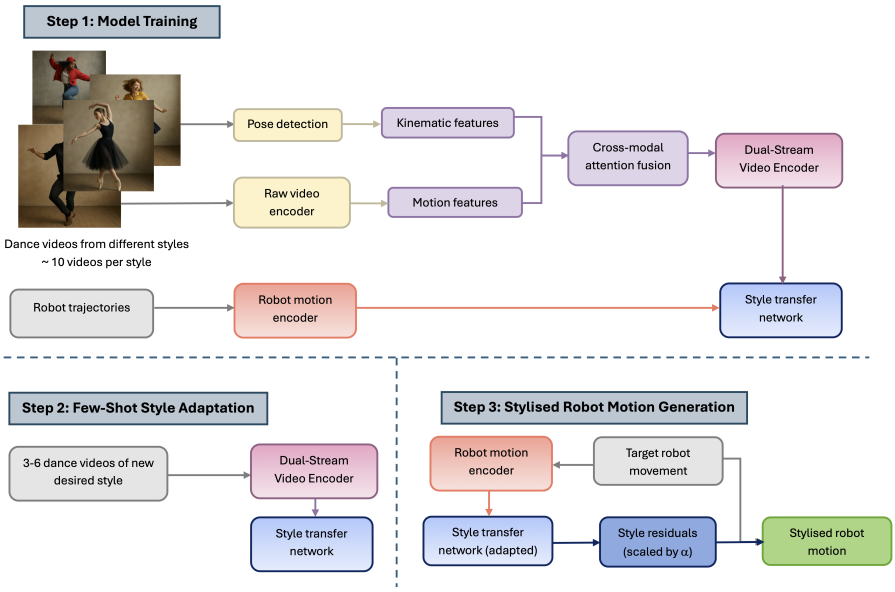

Figure 1: Schematic diagram summarising the proposed method.

2.1 Dual-Stream Video Encoder

Human movement style encompasses multiple complementary aspects requiring both visual motion patterns (overall flow, rhythm, aesthetic qualities) and precise joint kinematics (detailed movement mechanics and coordination patterns). Our dual-stream architecture processes both modalities in parallel, then fuses them through cross-modal attention to create a comprehensive style representation.

Raw Video Encoder: The raw video stream processes sequences of 128 frames at 64x64 resolution to extract motion features through temporal differentiation. We compute motion vectors as frame-to-frame differences to capture the temporal change of movements processed through convolutional layers with kernel size (8,8), stride (8,8) and 32 filters, followed by global average pooling to extract 32-dimensional features per timestep.

Pose Detection: Human pose data are extracted using MediaPipe (31) resulting in 33 features with 4 coordinates (x, y, z, visibility) per frame, organised into a 128x132 dimensional tensor. We compute pose velocities and accelerations to capture kinematic signatures, extracting movement magnitude, direction features, and joint coordination patterns. The final pose features combine rhythm characteristics with spatial coordination patterns, resulting in 10-dimensional temporal features.

Multi-Modal Fusion: The dual streams are combined through cross-modal attention that computes query-key relationships between raw video features and pose features. This bidirectional attention mechanism allows each stream to focus on complementary information from the other modality.

2.2 Robot Motion Encoder

The robot motion encoder transforms the 7 joint angle sequences of the Franka Panda over 128 time steps into a 512 dimensional representation that preserves essential kinematic characteristics while still allowing for style transfer. We employ a LSTM architecture with 128 hidden units in the first layer and 512 units in the second layer. This design progressively abstracts from low-level joint configurations to high-level motion patterns. This representation captures the overall trajectory of robot motion to be preserved during style transfer, ensuring that the robot maintains its intended movement while adopting stylistic characteristics.

2.3 Style Transfer Network

The style transfer network represents the core creative component where artistic expression is combined with robotic motion. The transformer-based architecture enables the system to learn complex

temporal relationships between style characteristics and motion modifications, while the residual learning approach ensures that style changes are additive rather than replacing the original motion entirely. The network combines the robot motions with style sequences through concatenation. The transformer architecture employs learned positional encoding using an embedding layer to maintain temporal relationships. Three temporal transformer layers with 8 attention heads, 256-dimensional model space, and 512-dimensional feed forward networks process the combined representation with 0.25 dropout rate. Each block applies multi-head self-attention followed by position-wise feed forward transformation with residual connections and layer normalisation. The network generates residuals followed by light temporal smoothing using 1D convolutions. The final styled motion blends 80% raw and 20% smoothed residuals.

2.4 Adversarial Training

Adversarial training ensures that generated robot motions remain physically plausible and natural looking. The loss function coordinates multiple objectives including adversarial realism, style expression, motion smoothness, and task preservation, with enhanced smoothness constraints specifically designed to prevent jittery or jerky motion that would be unsuitable for physical robot execution. The discriminator architecture processes 7-dimensional joint angle sequences through LSTM layers (128 and 64 hidden units) to capture temporal motion patterns, followed by dense layers (128 units with ReLU activation) for binary classification.

2.5 Few-Shot Style Adaptation

Real-world deployment requires the ability to quickly adapt to new movement styles. Few-shot adaptation addresses the constraint of extensive retraining and the requirement of collecting large datasets for each new style. By utilising the general motion understanding learned during initial training, the system can adapt to new styles using only 3-6 videos. The adaptation loss emphasises style consistency within the target domain while maintaining trajectory preservation.

2.6 Motion Generation and Creativity Control

Finally, motion generation represents the trained model applying learned style transformations to new movements to produce styled robot motions, with a creativity parameter providing control over the balance between style expression and original trajectory preservation. The transformer generates residual modifications that are scaled by a creativity parameter α:

$$\mathbf{q}_{styled} = \mathbf{q}_{original} + \alpha \mathbf{r}_{style} \tag{1}$$

where $\mathbf{q}_{styled}$ represents the styled motion and $\mathbf{r}_{style}$ are the generated style residuals. The creativity parameter α provides control over the balance between trajectory preservation and style expression.

3 Results and Evaluation

3.1 Methodology

We evaluated the system using 42 robot motion data files containing the 7 joint positions at 20Hz from "Robots, Dance, Different Bodies" involving professional dancers with different disabilities working with Franka Panda robot arms. The robot motion data represents authentic dance movements from collaborative sessions rather than pre-choreographed sequences. For style reference videos we used 41 publicly available demonstrations of different dance styles (ballet, jazz, flamenco, contemporary) and martial arts each between 10 and 20 minutes in length. We evaluate using four metrics: position accuracy (MAE between original and styled trajectories), acceleration change (magnitude of rhythmic modifications), style match score (alignment with reference style characteristics including rhythm intensity, acceleration patterns, movement coordination, and temporal dynamic) and overall style transfer score (cumulative strength of applied style residuals, indicating transformation magnitude). The overall style transfer score serves as a direct measure of how much creative transformation the system applies, making it particularly sensitive to creativity parameter adjustments.

Table 1: Style transfer performance across different styles and creativity levels.

Style	Creativity Parameter	Position Accuracy (MAE)	Acceleration Change	Style Match Score	Overall Style Transfer Score
Jazz	0.25	0.088	0.467	0.490	0.124
	0.50	0.184	1.165	0.500	0.329
	0.75	0.273	1.701	0.541	0.491
	1.00	0.352	3.111	0.500	0.899
Martial Arts	0.25	0.093	0.566	0.363	0.154
	0.50	0.174	1.561	0.380	0.438
	0.75	0.278	3.959	0.416	1.090
	1.00	0.355	3.555	0.423	1.042
Ballet	0.25	0.090	1.108	0.357	0.291
	0.50	0.172	2.828	0.403	0.779
	0.75	0.274	3.173	0.384	0.874
	1.00	0.366	2.185	0.353	0.641
Flamenco	0.25	0.091	0.532	0.380	0.143
	0.50	0.174	1.694	0.389	0.472
	0.75	0.260	2.135	0.394	0.625
	1.00	0.368	3.416	0.408	0.959
Contemporary	0.25	0.089	0.409	0.364	0.111
	0.50	0.186	1.058	0.369	0.293
	0.75	0.267	1.737	0.361	0.506
	1.00	0.353	2.933	0.368	0.821

3.2 Results and Discussion

Table 1 presents the evaluation metrics across five styles and four creativity parameter levels with the results demonstrating that the creativity parameter affects across all styles. Position accuracy consistently degrades as creativity increases (0.088-0.093 at 0.25 creativity vs 0.352-0.368 at 1.0 creativity), indicating the expected functional preservation limitation. Fig. 2 shows the trajectory of four joints across the applied styles with creativity=1. Meanwhile, acceleration changes increase dramatically with creativity (0.409-0.566 at low creativity vs 2.135-3.959 at high creativity), confirming that higher creativity parameters produce more dynamic, rhythmically distinct motions. Martial Arts exhibits the highest acceleration changes (3.959), consistent with the sharp movements in martial arts. Contemporary and Jazz show substantial but more moderate acceleration increases, reflecting their smoother yet dynamic characteristics. This variation in acceleration patterns across styles demonstrates successful capture and transfer of style specific movement qualities. However, the trajectory deviation with increased creativity introduces challenges for robot execution feasibility as while higher creativity settings achieve greater artistic expression, they may compromise the robot's ability to follow the original trajectory. This highlights the balance required between expressive movement and feasibility when selecting creativity levels.

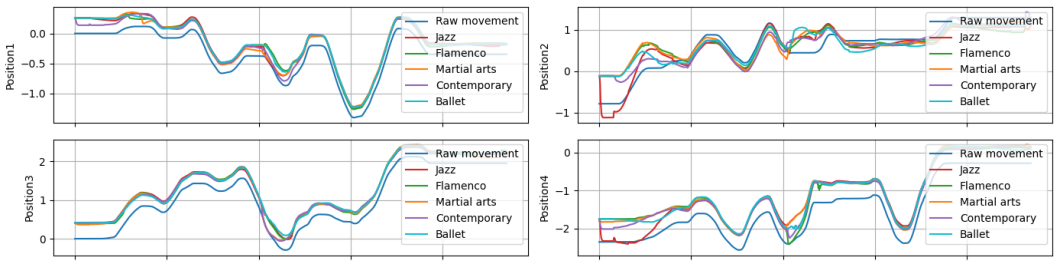

Figure 2: Comparison of raw and styled movements at creativity=1 for 4 of the robot joints.

Significant style differentiation has been achieved with each style exhibiting distinct characteristics as shown in Fig. 3. Jazz demonstrates the highest style match scores (0.490-0.541), suggesting alignment with reference jazz characteristics from the videos. However, style match scores show limited sensitivity to creativity parameter adjustments, with most styles exhibiting narrow score ranges compared to the substantial variations observed in position accuracy and acceleration change.

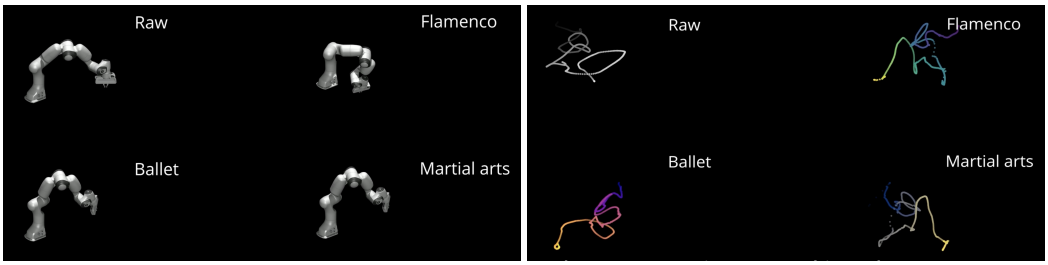

Figure 3: Visualisation of generated movements for original raw data and stylised Flamenco, Ballet and Martial arts motions a robot simulator (left) and the overall plotted trajectory (right).

This limited creativity responsiveness in style matching may reflect that the focus on high-level style characteristics reach saturation quickly or that kinematic constraints limit how closely human dance movements can be reproduced. Notably, different styles achieve optimal style match performance at different creativity levels with Jazz peaking at 0.75 creativity, Ballet at 0.5 creativity, and Flamenco showing steady improvement, reaching optimal performance at 1.0 creativity. This shows that each style has distinct optimal expression levels that may not simply align with the creativity parameter.

The Overall Style Transfer Score demonstrates the most pronounced creativity parameter sensitivity across all styles, providing clear evidence of style transformation capabilities. There is consistent increase in style transfer, ranging from 6.7x to 7.4x from minimum to maximum creativity with Martial Arts achieving the highest style transfer scores (1.042-1.090), reflecting the ability to apply the dynamic, sharp characteristic of martial arts movements. The consistent correlation between creativity and style transfer across all styles demonstrates the effectiveness of the creativity parameter. This shows that even when external style matching scores show limited variation, significant internal style transformations are occurring.

The computational efficiency and practical applicability of this approach is enhanced by the few-shot adaptation, requiring only 3-6 demonstration videos per style. The system was trained using an Nvidia GeForce RTX 5090 GPU with 64GB of RAM. This represents an advantage over traditional motion learning approaches that typically require hundreds or thousands of examples, making our system highly practical for real-world creative applications where extensive datasets are rarely available. However the dual-stream architecture introduces computational overhead that increases with video sequence length and number of pose keypoints.

While the proposed system demonstrates good performance and provides a new approach for the development of a collaborative performance between artists and robots, it also presents some limitations. The quality and type of movements present in the videos have a significant impact on the performance and creativity of the model and therefore the adapted movement. This is due to the model obtaining specific attributes associated with different styles from these videos, making quality video selection crucial. We also acknowledge potential concerns including the unauthorised use of proprietary creative content for training and the possibility of using this technology to create deceptively human-like robotic behaviours. Finally, our approach has only been evaluated with trajectory data from a single specific robot, a Franka Panda, which limits the evaluation of the proposed system.

4 Conclusion

We have presented a novel framework for few-shot cross-modal style transfer that enables robots to adopt human movement styles. Our dual-stream architecture with cross-modal attention successfully fuses visual motion and pose information, requiring only 3-6 demonstration videos to learn new styles. Our evaluation demonstrates that robots can serve as expressive partners in artistic contexts without sacrificing functional performance, creating new possibilities for human-robot collaboration. While our computational metrics demonstrate successful style differentiation, evaluating creative value requires human perception studies beyond this initial technical demonstration (32). Future work must evaluate whether dancers and audiences perceive the generated motions as artistically meaningful contributions to collaborative performance as well as exploring more comprehensive evaluation of generated motion feasibility including physical robot trials.

Acknowledgments and Disclosure of Funding

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) through the Turing AI World Leading Researcher Fellowship: Somabotics: Creatively Embodying Artificial Intelligence [grant number EP/Z534808/1]. We would like to extend our heartfelt thanks to our friends and colleagues who participated in the research project Embodied Trust in TAS: Robots, Dance, Diferent Bodies; Sarah Whatley, Kate Marsh, Welly O'Brien, and Kimberly Harvey. We would also like to acknowledge our creative partners, the Candoco Dance Company. The control software for the robotic arms was developed with foundational work contributed by Dr. Joseph Bolarinwa.

References

[1] J. Rezwana and M. L. Maher, "Designing creative AI partners with COFI: A framework for modeling interaction in human-AI co-creative systems," *ACM Transactions on Computer-Human Interaction*, vol. 30, no. 5, pp. 1–28, 2023.

[2] S. Benford, E. Schneiders, J. P. M. Avila, P. Caleb-Solly, P. R. Brundell, S. Castle-Green, F. Zhou, R. Garrett, K. Höök, S. Whatley *et al.*, "Somatic safety: An embodied approach towards safe human-robot interaction," in *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2025, pp. 429–438.

[3] R. Garrett, P. Brundell, S. Castle-Green, K. Hawkins, P. Tennent, F. Zhou, A. Lampinen, K. Höök, and S. Benford, "Friction in processual ethics: Reconfiguring ethical relations in interdisciplinary research," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025, pp. 1–15.

[4] R. Garrett, K. Hawkins, P. Brundell, S. Castle-Green, P. Tennent, F. Zhou, A. Lampinen, K. Höök, and S. Benford, "In the moment of glitch: Engaging with misalignments in ethical practice," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025, pp. 1–18.

[5] J.-W. Hong and N. M. Curran, "Artificial intelligence, artists, and art: attitudes toward artwork produced by humans vs. artificial intelligence," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 15, no. 2s, pp. 1–16, 2019.

[6] E. Cetinic and J. She, "Understanding and creating art with AI: Review and outlook," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 18, no. 2, pp. 1–22, 2022.

[7] Z. Epstein, A. Hertzmann, I. of Human Creativity, M. Akten, H. Farid, J. Fjeld, M. R. Frank, M. Groh, L. Herman, N. Leach *et al.*, "Art and the science of generative AI," *Science*, vol. 380, no. 6650, pp. 1110–1111, 2023.

[8] H.-K. Ko, G. Park, H. Jeon, J. Jo, J. Kim, and J. Seo, "Large-scale text-to-image generation models for visual artists' creative works," in *Proceedings of the 28th international conference on intelligent user interfaces*, 2023, pp. 919–933.

[9] J. Huang, "The art of AI: A human-centered AI (HCAI) user study of integrating image-generative tools in visual art workflows: The case of Adobe Firefly," 2024.

[10] J. Oppenlaender, H. Johnston, J. M. Silvennoinen, and H. Barranha, "Artworks reimagined: Exploring human-AI co-creation through body prompting," *Proceedings of the ACM on Human-Computer Interaction*, vol. 9, no. 4, pp. 1–34, 2025.

[11] A. Ferracani, S. Ricci, F. Principi, G. Becchi, N. Biondi, A. Del Bimbo, M. Bertini, and P. Pala, "An AI-powered multimodal interaction system for engaging with digital art: A human-centered approach to HCI," in *International Conference on Human-Computer Interaction*. Springer, 2025, pp. 281–294.

[12] J. Rezwana and M. L. Maher, "Designing creative ai partners with cofi: A framework for modeling interaction in human-ai co-creative systems," *ACM Transactions on Computer-Human Interaction*, vol. 30, no. 5, pp. 1–28, 2023.

[13] Y. Fu, M. Newman, L. Going, Q. Feng, and J. H. Lee, "Exploring the collaborative co-creation process with AI: A case study in novice music production," in *Proceedings of the 2025 ACM Designing Interactive Systems Conference*, 2025, pp. 1298–1312.

[14] L. Sun, P. Chen, W. Xiang, P. Chen, W.-y. Gao, and K.-j. Zhang, "Smartpaint: a co-creative drawing system based on generative adversarial networks," *Frontiers of Information Technology & Electronic Engineering*, vol. 20, no. 12, pp. 1644–1656, 2019.

[15] A.-S. Maerten and D. Soydaner, "From paintbrush to pixel: A review of deep neural networks in AI-generated art," *arXiv preprint arXiv:2302.10913*, 2023.

[16] J. P. Ferreira, T. M. Coutinho, T. L. Gomes, J. F. Neto, R. Azevedo, R. Martins, and E. R. Nascimento, "Learning to dance: A graph convolutional adversarial network to generate realistic dance motions from audio," *Computers & Graphics*, vol. 94, pp. 11–21, 2021.

[17] M. Trajkova, D. Long, M. Deshpande, A. Knowlton, and B. Magerko, "Exploring collaborative movement improvisation towards the design of LuminAI—a co-creative AI dance partner," in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–22.

[18] T. Wang, W. Q. Toh, H. Zhang, X. Sui, S. Li, Y. Liu, and W. Jing, "Robocodraw: Robotic avatar drawing with gan-based style transfer and time-efficient path optimization," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 06, 2020, pp. 10 402–10 409.

[19] M. Tang, "Human and machine symbiosis-an experiment of human and robot co-creation of calligraphy-style drawing," in *International Conference on Human-Computer Interaction*. Springer, 2022, pp. 462–467.

[20] G. Venture and D. Kulić, "Robot expressive motions: a survey of generation and evaluation methods," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 8, no. 4, pp. 1–17, 2019.

[21] C. Gomez Cubero, M. Pekarik, V. Rizzo, and E. Jochum, "The robot is present: Creative approaches for artistic expression with robots," *Frontiers in Robotics and AI*, vol. 8, p. 662249, 2021.

[22] A. LaViers and M. Egerstedt, "Style based robotic motion," in *2012 American Control Conference (ACC)*. IEEE, 2012, pp. 4327–4332.

[23] M. Abdul-Massih, I. Yoo, and B. Benes, "Motion style retargeting to characters with different morphologies," in *Computer Graphics Forum*, vol. 36, no. 6. Wiley Online Library, 2017, pp. 86–99.

[24] K. Aberman, Y. Weng, D. Lischinski, D. Cohen-Or, and B. Chen, "Unpaired motion style transfer from video to animation," *ACM Transactions on Graphics (TOG)*, vol. 39, no. 4, pp. 64–1, 2020.

[25] S. M. A. Akber, S. N. Kazmi, S. M. Mohsin, and A. Szczęsna, "Deep learning-based motion style transfer tools, techniques and future challenges," *Sensors*, vol. 23, no. 5, p. 2597, 2023.

[26] T. Zhang and H. Tang, "Style transfer: A decade survey," *arXiv preprint arXiv:2506.19278*, 2025.

[27] J. Young, K. Ishii, T. Igarashi, and E. Sharlin, "Style by demonstration: teaching interactive movement style to robots," in *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*, 2012, pp. 41–50.

[28] X. Zhu, Z. Chen, and J. Chen, "Stylized table tennis robots skill learning with incomplete human demonstrations," *arXiv preprint arXiv:2309.08904*, 2023.

[29] S. Adams, T. Cody, and P. A. Beling, "A survey of inverse reinforcement learning," *Artificial Intelligence Review*, vol. 55, no. 6, pp. 4307–4346, 2022.

[30] Y. Yuan, Z. Li, T. Zhao, and D. Gan, "DMP-based motion generation for a walking exoskeleton robot using reinforcement learning," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 5, pp. 3830–3839, 2020.

[31] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee, W.-T. Chang, W. Hua, M. Georg, and M. Grundmann, "MediaPipe: A framework for building perception pipelines," 2019. [Online]. Available: https://arxiv.org/abs/1906.08172

[32] A. Newell, J. C. Shaw, and H. A. Simon, "The processes of creative thinking." *Contemporary approaches to creative thinking: A symposium held at the University of Colorado.*, pp. 63–119, 7 1962.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: All claims related to the developed style transfer system and its performance made in the abstract and introduction are reflected in the paper's contributions in Section 3.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The limitations presented by the proposed system and the data used for its evaluation can be found within Section 2.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [NA]

   Justification: The paper includes no theoretical results only experimental results.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: Section 2 describes the system architecture used in this paper. Also, all code has been released publicly on GitHub and referenced in Section 1.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: All code has been released publicly on GitHub and referenced in Section 1.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [Yes]

   Justification: The experimental setup is described but as this work is focused on novel generation there is no test train split instead all data is used for training and then new motions are generated and evaluated using the metrics described in Section 3.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [No]

   Justification: This evaluation serves as an initial proof-of-concept demonstration of the system's capabilities. The deterministic evaluation on our dataset provides baseline performance metrics, though statistical significance testing across multiple runs would strengthen future evaluations. Given the creative and subjective nature of dance style transfer, this evaluation focuses on demonstrating the system's range of capabilities across different styles and creativity parameters using our dataset of authentic dance collaborations.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: The compute resources required is discussed in Section 3.

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The paper follows all NeurIPS code of Ethics.

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: The potential impact and misuse of the work is discussed in Section 3 and Section 4.

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

    Justification: We consider that, while the proposed model and application could present some copyright concerns regarding the ownership of the produced output, we consider that it does not have a high risk of misuse.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: The system was evaluated using publicly available movement video demonstrations and our own dataset of robot movements. While the videos informed our evaluation methodology, the specific choreographic content is not reproduced or distributed, and the system generates novel robot movements rather than copying human dance sequences. The dancers who collected the robot motion data are credited within the paper (although removed for initial submission due to anonymisation guidelines). However, the specific dancers and choreographers from the videos are not individually credited within the paper.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: The associated code has been publicly released in a GitHub repository and referenced in Section 1.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: this work does not use crowdsourcing experiments and research with human subjects.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

    Answer: [NA]

    Justification: The method proposed in this application does not involve or make use in any way of LLMs.