1. Ahn, H., Kim, J., Kim, K. & Oh, S. Generative autoregressive networks for 3d dancing move synthesis from music. IEEE Robotics And Automation Letters. 5, 3501-3508 (2020)

2. Aldausari, N., Sowmya, A., Marcus, N. & Mohammadi, G. Video generative adversarial networks: a review. ACM Computing Surveys (CSUR). pp. 1-25 (2022).

3. Aldausari, N., Sowmya, A., Marcus, N. & Mohammadi, G. Diverse Audio-to-Video GAN using Multiscale Image Fusion. Australasian Joint Conference On Artificial Intelligence. pp. 29-42 (2022). Springer.

4. Aldausari, N., Sowmya, A., Marcus, N. & Mohammadi, G. Cascaded siamese self-supervised audio to video GAN. Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition. pp. 4691-4700 (2022).

5. Alemi, O., Françoise, J. & Pasquier, P. GrooveNet: Real-time music-driven dance movement generation using artificial neural networks. Networks. 8, 26 (2017)

6. Chan, C., Ginosar, S., Zhou, T. & Efros, A. Everybody dance now. Proceedings Of The IEEE/CVF International Conference On Computer Vision. pp. 5933-5942 (2019)

7. Chen, L., Cao, C., Torre, F., Saragih, J., Xu, C. & Sheikh, Y. High-fidelity face tracking for ar/vr via deep lighting adaptation. Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition. pp. 13059-13069 (2021)

8. Dhariwal, P. & Nichol, A. Diffusion models beat gans on image synthesis. Advances In Neural Information Processing Systems. 34 pp. 8780-8794 (2021)

Rhythm Fusion 15

9. Di, S., Jiang, Z., Liu, S., Wang, Z., Zhu, L., He, Z., Liu, H. & Yan, S. Video background music generation with controllable music transformer. Proceedings Of The 29th ACM International Conference On Multimedia. pp. 2037-2045 (2021)

10. Doukas, M., Zafeiriou, S. & Sharmanska, V. Headgan: One-shot neural head syn-

thesis and editing. Proceedings Of The IEEE/CVF International Conference On Computer Vision. pp. 14398-14407 (2021)

11. Duan, Y., Shi, T., Zou, Z., Qin, J., Zhao, Y., Yuan, Y., Hou, J., Wen, X. & Fan, C. Semi-supervised learning for in-game expert-level music-to-dance translation. ArXiv Preprint ArXiv:2009.12763. (2020)

12. Foundation, B. Blender (software). , https://www.blender.org/ Available: https://www.blender.org/

13. Goodfellow, I., Pouget-Abadie, J. & Others Generative adversarial nets. Advances In Neural Information Processing Systems. pp. 2672-2680

14. Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. Advances In Neural Information Processing Systems. 33 pp. 6840-6851 (2020)

15. Holden, D., Saito, J. & Komura, T. A deep learning framework for character motion synthesis and editing. ACM Transactions On Graphics (TOG). 35, 1-11 (2016)

16. Huang, R., Hu, H., Wu, W., Sawada, K., Zhang, M. & Jiang, D. Dance revolution: Long-term dance generation with music via curriculum learning. ArXiv Preprint ArXiv:2006.06119. (2020)

17. Huang, P., Yang, F. & Wang, Y. Learning identity-invariant motion representations for cross-id face reenactment. Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition. pp. 7084-7092 (2020)

18. Kim, H., Elgharib, M., Zollhöfer, M., Seidel, H., Beeler, T., Richardt, C. & Theobalt, C. Neural style-preserving visual dubbing. ACM Transactions On Graphics (TOG). 38, 1-13 (2019)

19. Kingma, D. & Welling, M. Auto-encoding variational bayes. ArXiv Preprint ArXiv:1312.6114. (2013)

20. KR, P., Mukhopadhyay, R., Philip, J., Jha, A., Namboodiri, V. & Jawahar, C. Towards automatic face-to-face translation. Proceedings Of The 27th ACM Interna-

tional Conference On Multimedia. pp. 1428-1436 (2019)

21. Lee, H., Yang, X., Liu, M., Wang, T., Lu, Y., Yang, M. & Kautz, J. Dancing to music. Advances In Neural Information Processing Systems. 32 (2019)

22. Li, B., Zhao, Y., Zhelun, S. & Sheng, L. Danceformer: Music conditioned 3d dance generation with parametric motion transformer. Proceedings Of The AAAI Conference On Artificial Intelligence. 36, 1272-1279 (2022)

23. Li, R., Yang, S., Ross, D. & Kanazawa, A. Ai choreographer: Music conditioned 3d dance generation with aist++. Proceedings Of The IEEE/CVF International Conference On Computer Vision. pp. 13401-13412 (2021)

24. Li, J., Yin, Y., Chu, H., Zhou, Y., Wang, T., Fidler, S. & Li, H. Learning to generate diverse dance motions with transformer. ArXiv Preprint ArXiv:2008.08171. (2020)

25. Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., & Black, M. J. SMPL: A skinned multi-person linear model. Seminal Graphics Papers: Pushing the Boundaries, Volume 2, pp. 851-866 (2023).

26. Müller, M., Röder, T. & Clausen, M. Efficient content-based retrieval of motion capture data. ACM SIGGRAPH 2005 Papers. pp. 677-685 (2005)

27. Oord, A., Dieleman, S., Zen, H. & Others Wavenet: A generative model for raw audio. ArXiv Preprint ArXiv:1609.03499. (2016)

28. Onuma, K., Faloutsos, C. & Hodgins, J. FMDistance: A Fast and Effective Distance Function for Motion Capture Data.. Eurographics (Short Papers). pp. 83-86 (2008)

29. Prajwal, K. R., Mukhopadhyay, R., Namboodiri, V. P. & Jawahar, C. V. A lip sync expert is all you need for speech to lip generation in the wild. Proceedings Of The 28th ACM International Conference On Multimedia. pp. 484-492 (2020).

30. Rezende, D. & Mohamed, S. Variational inference with normalizing flows. International Conference On Machine Learning. pp. 1530-1538 (2015)

31. Siyao, L., Yu, W., Gu, T., Lin, C., Wang, Q., Qian, C., Loy, C. & Liu, Z. Bailando:

3d dance generation by actor-critic gpt with choreographic memory. Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition. pp. 11050-11059 (2022)

32. Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N. & Ganguli, S. Deep unsupervised learning using nonequilibrium thermodynamics. International Conference On Machine Learning. pp. 2256-2265 (2015)

33. Song, Y. & Ermon, S. Generative modeling by estimating gradients of the data distribution. Advances In Neural Information Processing Systems. 32 (2019)

34. Sun, G., Wong, Y., Cheng, Z., Kankanhalli, M., Geng, W. & Li, X. DeepDance: music-to-dance motion choreography with adversarial learning. IEEE Transactions On Multimedia. 23 pp. 497-509 (2020)

35. Tang, T., Jia, J. & Mao, H. Dance with melody: An lstm-autoencoder approach to music-oriented dance synthesis. Proceedings Of The 26th ACM International Conference On Multimedia. pp. 1598-1606 (2018)

36. Tang, T., Jia, J. & Mao, H. Dance with melody: An lstm autoencoder approach to music oriented dance synthesis. Proceedings Of The 26th ACM International Conference On Multimedia. pp. 1598-1606 (2018)

37. Tevet, G., Raab, S., Gordon, B., Shafir, Y., Cohen-Or, D. & Bermano, A. Human motion diffusion model. ArXiv Preprint ArXiv:2209.14916. (2022)

38. Tseng, J., Castellon, R. & Liu, K. Edge: Editable dance generation from music. Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition. pp. 448-458 (2023)

39. Wu, W., Zhang, Y., Li, C., Qian, C. & Loy, C. Reenactgan: Learning to reenact faces via boundary transfer. Proceedings Of The European Conference On Computer Vision (ECCV). pp. 603-619 (2018)

40. Yang, S., Yang, Z. & Wang, Z. LongDanceDiff: Long-term Dance Generation with

Conditional Diffusion Model. ArXiv Preprint ArXiv:2308.11945. (2023)

41. Yang, Z., Zhu, W. & Others TransMoMo: Invariance-Driven Unsupervised Video Motion Retargeting. Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition. pp. 5306-5315

42. Zakharov, E., Shysheya, A., Burkov, E. & Lempitsky, V. Few-shot adversarial learning of realistic neural talking head models. Proceedings Of The IEEE/CVF International Conference On Computer Vision. pp. 9459-9468 (2019)

43. Zhou, Y., Wang, Z., Fang, C., Bui, T. & Berg, T. Dance dance generation: Motion transfer for internet videos. Proceedings Of The IEEE/CVF International Conference On Computer Vision Workshops. pp. 0-0 (2019)

44. Zhou, H., Liu, J., Liu, Z., Liu, Y. & Wang, X. Rotate-and-render: Unsupervised photorealistic face rotation from single-view images. Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition. pp. 5911-5920 (2020)

45. Zhang, J., Zeng, X., Wang, M., Pan, Y., Liu, L., Liu, Y., Ding, Y. & Fan, C. Freenet: Multi-identity face reenactment. Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition. pp. 5326-5335 (2020)

46. Zhuang, W., Wang, C. & Others Music2dance: Music-driven dance generation using wavenet. ArXiv Preprint ArXiv:2002.03761. (2020)