#79: Intelligent body tracking and agent in Virtual reality to study embodiment

Authors:
Amanda Huang, Jason Xavier
Supervisor:
Danielle Lottridge
Co-supervisor:
Elliott Wen

Background

This project belongs to an overarching project where users can dance in a VR environment and view themselves as abstract avatars. The stakeholders include dancers from the University of Auckland's dance faculty. Initially, the project aimed to incorporate a 3D AI dance model into the existing framework. However, after finding the quality of several available models inadequate, we shifted our focus to analysing quantitative metrics from recorded dance motions and exploring differences based on their clusters.

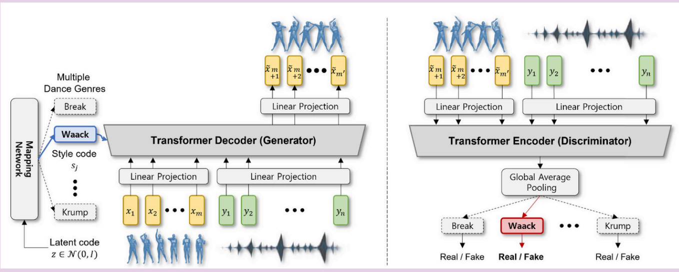
Objectives

- Investigate the feasability of integrating an existing 3D AI dance model into the VR project.
- Analyse the quantitative differences in dance motions across various clusters.

AI Models Overview

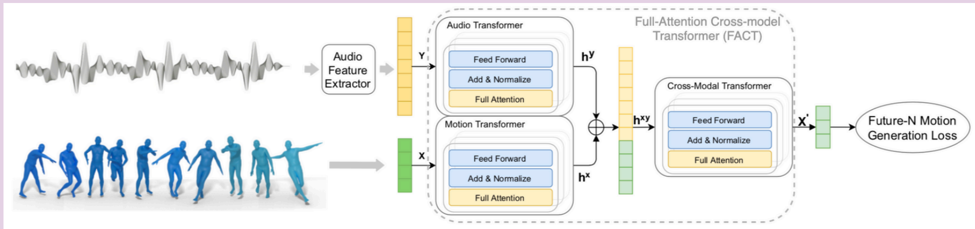
We explored two AI models capable of generating dance motions based on input motions and music.

MNET [1]



- Generative adversarial network with transformer architecture
- Modified configurations and resolved bugs to get the model in a runnable state
- Trained the model from scratch
- Modified a visualisation tool to display the input motions in a human-friendly manner

FACT [2]



- Full-attention cross-modal transformer network
- Modified the code to accept custom motion and audio inputs
- Modified the output motion format to ensure compatibility with an existing visualisation tool

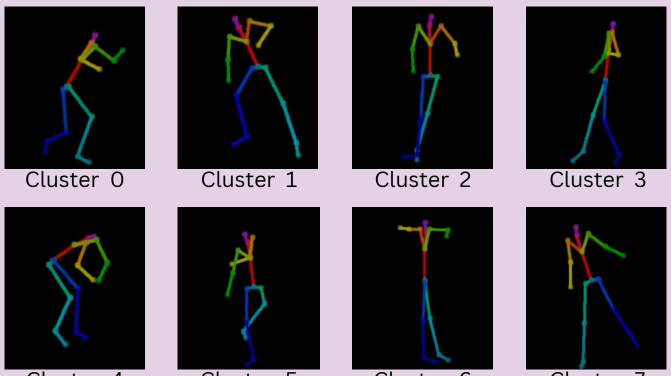
Both models were found to be inadequate, as they struggled to handle data outside of their training sets well. After discussions with supervisors and stakeholders, it was determined that integrating these models into the project was unviable, leading us to shift our focus to motion analytics instead.

Motion Analytics

Context

Recorded dance motions were categorised into eight clusters using k-means clustering based on Laban metrics. The goal of our motion analytics is to determine how quantitative metrics differentiate clusters from each other.

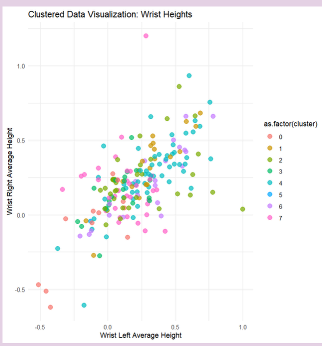
Clusters

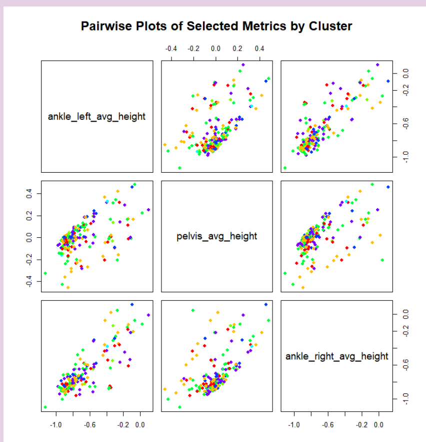


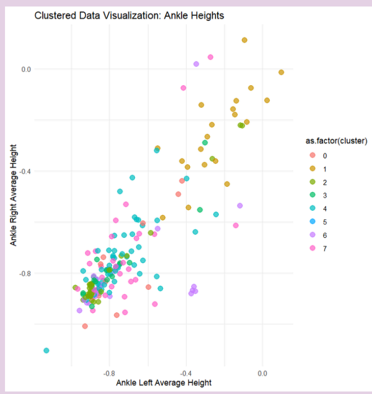
Cluster 0 Cluster 1 Cluster 2 Cluster 3
Cluster 4 Cluster 5 Cluster 6 Cluster 7

Methodology

- We worked with motions converted from the standard SMPL format [3] to a custom format based on x, y, z coordinates instead of joint rotations.
- For each motion, we generated overall quantitative metrics, including max, min, average, and standard deviation, across parameters such as height, distance, velocity, and acceleration. These metrics were analysed for key body points: left ankle, right ankle, pelvis, solar plexus, left wrist, and right wrist.

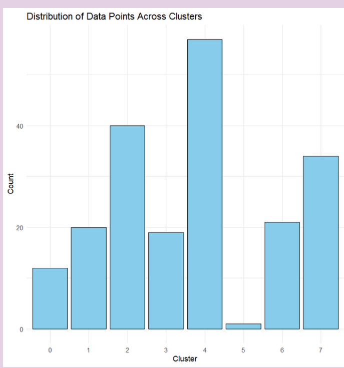






Results

- Observed a strong association between complementary joints (e.g the left and right ankles and wrists)
- Variation in the key metrics was able to be visually estimated through our graphs



Future Work

- Investigate alternative data visualisation methods
- Investigate the differences in quantitative metrics between the input and output motions generated by the AI model

UNIVERSITY OF AUCKLAND
Waipapa Taumata Rau
NEW ZEALAND

ENGINEERING
DEPARTMENT OF ELECTRICAL, COMPUTER, AND SOFTWARE ENGINEERING

References
[1] J.-W. Kim, H. Oh, S. Kim, H. Tong, and S. Lee, "A Brand New Dance Partner: Music-Conditioned Pluralistic Dancing Controlled by Multiple Dance Genres," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2022, doi: https://doi.org/10.1109/cvpr52688.2022.00348.
[2] R. Li, S. Yang, D. Ross, and A. Kanazawa, "AI Choreographer: Music Conditioned 3D Dance Generation with AIST++," Oct. 2021, doi: https://doi.org/10.1109/iccv48922.2021.01315.
[3] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black, "SMPL," *ACM Transactions on Graphics*, vol. 34, no. 6, pp. 1–16, Nov. 2015, doi: https://doi.org/10.1145/2816795.2818013.