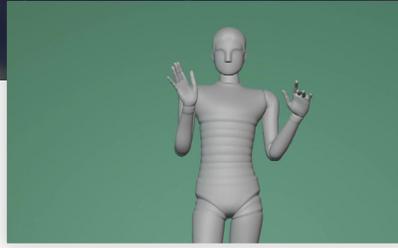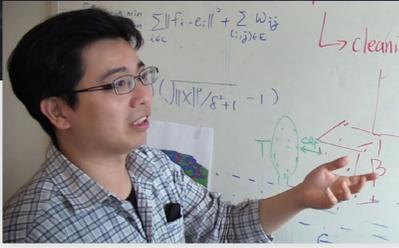# 3D Gesture Performance Synthesis

2021 - Current
Project Leader: Andrew Feng

Background

Creating a believable virtual human that can interact with human users, and mimic real-world face-to-face communications, requires extensive research. Previously such research depended on traditional workflow such as manual creations or motion capture, which are expensive and difficult to scale. As a result, the existing gesture data are limited in duration and/or animation quality, making them less ideal for training gesture synthesis models using machine learning methods. This project for 3D Gesture Performance Synthesis enables generation of multi-modal gesture datasets, with 3D gesture motions and corresponding speech audios extracted from in-the-wild monocular videos such as TED Talk video data. The new dataset also enables us to develop a personalized gesture synthesis model that reproduces the gesturing styles of individual speakers.

Objectives

Motivated by recent progress in human mesh recovery (HMR), ICT has now developed a tool for extracting avatar-ready gesture motions from monocular videos with improved animation quality. The tool utilizes a variational autoencoder (VAE) to refine raw gesture motions. The resulting gestures are in a unified pose representation that includes both body and finger motions and can be readily applied to a virtual avatar via online motion retargeting. ICT has validated the proposed tool on existing datasets and created the refined dataset TED-SMPLX by re-processing videos from the original TED dataset. Utilizing the gesture dataset, we have developed an ML-based gesture synthesis system that can generate novel 3D gesture animations from speech input.

Results

This project is on-going and documented in peer-reviewed academic papers and publications including: Association for Computing Machinery https://doi.org/10.1145/3561975.3562953

Next Steps

While the gesture synthesis model is already able to produce novel gestures, we plan to further enhance its quality

and controllability by allowing the users to adjust the styles of synthesized motions. Our key motivation is to utilize the recent advances in generative AI to develop a new motion synthesis framework. This allows both scalable and controllable human motion creations that match the user intent.

Published academic research papers are available from https://ict.usc.edu/research/publications
(Search engine keyword: USC ICT Publications)

Project Leader: Andrew Feng

Established in 1999, the USC Institute for Creative Technologies (ICT) is a Department of Defense (DoD) University Affiliated Research Center (UARC), sponsored by the US Army. Harnessing Hollywood-derived creativity with academic innovation and military-domain expertise, ICT conducts award-winning R&D in Artificial Intelligence (AI), Computer Graphics, Geospatial Sciences, Human Performance, Learning Sciences, Modeling, Simulation & Gaming, Mixed Reality (MxR), Medical VR, Narrative, and Virtual Humans.