



Caoyuan Ma

Wuhan University

Computer Science and Technology

+86 13146496689

✉ macaoyuan@whu.edu.cn

🐙 GitHub: Miles629

🌐 <https://miles629.github.io>

EDUCATION

- **Wuhan University - School of Computer and Science** 2022.9-now
Computer Science and Technology - Human-centric AIGC & 3D Master
- **Wuhan University - School of Computer and Science** 2018.9-2022.7
Software Engineering - GPA 3.87/4.0, Top 1% in 2 academic years Undergraduate

EXPERIENCE

- **JD Explore Academy** 2022.09-2023.10
The current method of HumanNeRF is mainly designed to synthesize images of arbitrary viewpoints through learning 4D representation. This design has a large number of artifacts when synthesizing images with new human poses, and requires a lot of training data to fit the human body, which limits the application of existing methods. We have designed a new approach which can learn high-quality human representation in limited data and can synthesize artifact-free human body images with any poses.

RESEARCH

- **A Simple yet Effective Approach to Animate HumanNeRF with Diverse Poses** 1st author CVPR24
Bg: Current methods require many training data, and show artifacts in novel pose. Accepted
Method: We believe that explicit prior human information can be used to guide the selection of sampling points for learning implicit representations. The highly-coupled explicit prior and implicit representation are achieved through local convolution filters.

Results: Compared with the state-of-the-art HumanNeRF method, we have reduced about **95%** of the training time and increased the inference speed by **15** times while using only **1%** of the training parameters. We have also achieved higher quality results in synthesizing images of new poses.

- **TAAT: Think and Act from Arbitrary Texts** 1st author MM24
Bg: Current methods only generate motions from texts with explicit actions. Submitted
Method: We leverage the emergent capabilities of LLMs to analyze arbitrary text and extract explicit or implicit actions. We generate a continuous and controllable motion sequences by constructing a stack of actions obtained from the text analysis.

Results: We have implemented the first Text2Motion method that can generate actions from scene text containing potential actions. This extends the applicability of the task to unconstrained scenario interactions.

One 1st-author AVSS, two co-1st-author papers under review. Contributed to one IST and one NIPS and one AAAI under review.

AWARDS

- **Scholarship:** National Scholarship * 2 (2/208), First Class Scholarship * 2 (Top 5%)
- **Honor:** Outstanding graduates of Wuhan University (Top 10%)
- **Competition:** Meritorious Winner of Mathematical Contest in Modeling(MCM)
First prize of National Undergraduate IoT Contest.