Data-Driven 3D Human Motion Modeling
for Reconstruction and Conditional Generation

Name: Huakun Liu

Laboratory: Cybernetics and Reality Engineering Laboratory

Supervisor: Prof. Kiyoshi Kiyokawa

Abstract: Human motion is one of the most important signals through which people act, communicate, and interact with the physical and social world. Computational models of 3D human motion, therefore, provide a foundation for many human-centered technologies, including embodied artificial intelligence, extended reality, and human-computer interaction. This thesis studies data-driven 3D human motion modeling for two closely related problems: robust reconstruction and conditional generation of motions. The first part of the thesis focuses on reconstruction. Instead of assuming dense optical motion-capture systems, it addresses a light setting in which full-body 3D motion is recovered from sparse, noisy, and partially unreliable wearable sensing signals. The proposed methods combine learned motion priors with inertial sensing, ultra-wideband ranging, and uncertainty estimation to improve body pose and trajectory estimation in the presence of ambiguity, drift, and sensor noise. The second part of the thesis investigates conditional generation. This thesis investigates how motion can be synthesized not only as plausible kinematic sequences but also as behavior that responds to semantic, affective, and social conditions. It studies the generation of emotional body motion using culturally grounded motion-capture data, analyzes how generated expressions are perceived by human observers, and examines whether large language models can serve as scalable proxies for estimating the distribution of human emotional perception. It further extends generation from isolated motion clips to continuous streams of motions that include both solo and interactive behavior. Overall, this thesis shows how learned motion models can transform ambiguous sensor observations into reliable 3D motion estimates and transform high-level conditions into motion that is physically plausible and perceptually meaningful. These contributions support more accurate, accessible, and controllable human-motion technologies for interactive virtual and embodied systems.