The Hong Kong University of Science and Technology

Department of Mathematics

PhD THESIS EXAMINATION

Animatable 3D Human Avatar Modeling from Multi-view Images

By

Mr. Zhiyuan YU

ABSTRACT

Animatable 3D human avatar modeling is a long-standing research problem in computer vision and computer graphics. This task aims to faithfully capture and synthesize realistic human appearances and movements from visual observations, thus enabling immersive and interactive experiences in virtual reality, augmented reality, telepresence, film production, and gaming. Despite significant advancements, current approaches still face critical challenges, including limited capability in modeling dynamic textures, difficulties in real-time rendering of animatable avatars, and poor generalization to high-quality 3D human reconstruction and animation. This thesis addresses these challenges by proposing a series of interconnected techniques, paving the way for more accessible, efficient, and realistic digital human content creation. Specifically, our contributions are threefold:

In the first part, we present 4K4D++, a dynamic human reconstruction method leveraging a temporally continuous 3D Gaussian representation. By explicitly modeling temporal coherence and utilizing a lightweight image-based rendering appearance module, 4K4D++ significantly enhances reconstruction efficiency and captures fine-grained dynamic details, surpassing state-of-the-art methods in rendering quality and speed. Building upon this enhanced dynamic reconstruction framework, the second part of the thesis presents NBAvatar, a human avatar reconstruction approach that enables both high-fidelity appearance rendering and real-time animation. Our key innovation lies in neural blendfeatures, a pose-dependent feature representation, substantially improving texture fidelity and reducing computational overhead. Extensive evaluations on public datasets demonstrate that NBAvatar achieves comparable visual fidelity to state-of-the-art methods while providing significant speedups in both training (up to 2.5×) and animation (up to 3.1×). Despite these advancements, both 4K4D++ and NBAvatar rely heavily on dense multi-view inputs and iterative optimization procedures, limiting their practicality for everyday consumer use. To overcome these limitations, the third part of the thesis proposes HumanRAM, a generalizable, unified feed-forward framework for human reconstruction and animation from single-view or sparse-view inputs. HumanRAM integrates parametric human priors into the large reconstruction model, enabling high-quality human reconstruction and high-fidelity human animation. Comprehensive experiments validate that our approach significantly outperforms existing methods in novel view and novel pose synthesis, demonstrating robust generalization capabilities on real-world data.

Date : **5 August 2025, Tuesday**

Time : **2:00 pm**

Venue : **Room 4472 (Lifts 25/26)**

***Thesis Examination Committee*:**

Chairman	**:**	**Prof. Yukinori HIRANO, LIFS/HKUST**
Thesis Supervisor	**:**	**Prof. Can YANG, MATH/HKUST**
Member	**:**	**Prof. Dong XIA, MATH/HKUST**
Member	**:**	**Prof. Xinzhou GUO, MATH/HKUST**
Member	**:**	**Prof. Yuan LIU, IEDA/HKUST**
External Examiner	**:**	**Prof. Ruizhen HU,** **Computer Science & Software Engineering/ Shenzhen University**

(Open to all faculty and students)

The student's thesis is now being displayed on the reception counter in the General Administration Office (Room 3461).