Move Your Twin: Live Digital Human Animation

Giulia Martinelli

DISI University of Trento, Trento, Italy
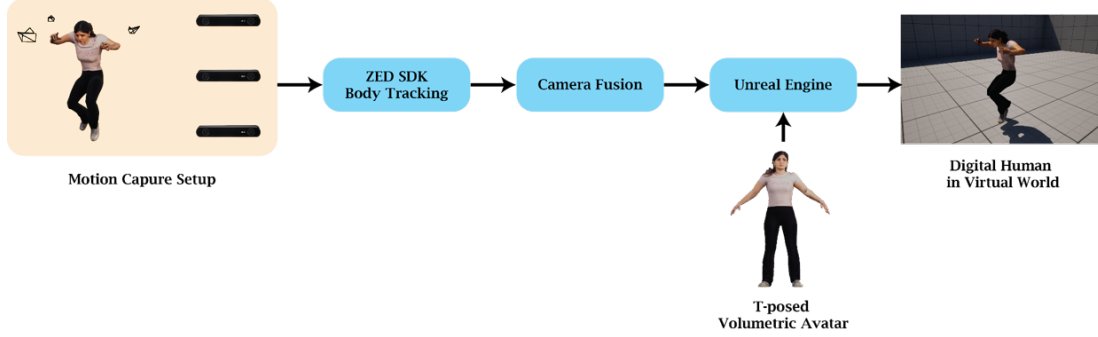
giulia.martinelli-2@unitn.it

Fig. 1: **Pipeline of the proposed demo.** A three-camera ZED 2i setup captures the user through real-time markerless body tracking. The multi-view skeletal data is fused and streamed to Unreal Engine, where different digital characters are animated live based on the user's movements.

Abstract—Digital Humans are becoming central in virtual production, XR, sports analysis, and interactive applications, where convincing appearance and motion are essential. Traditional optical motion-capture systems offer high accuracy but require markers, controlled environments, and complex calibration. Markerless approaches provide a more accessible alternative. This demo showcases a real-time markerless motion-capture pipeline using three synchronized ZED 2i stereo cameras to estimate full-body pose and stream it directly into Unreal Engine. Users can animate different digital characters, experiencing first-hand the full pipeline from human motion to digital animation.

I. INTRODUCTION

Creating realistic Digital Humans requires addressing multiple challenges, including skeleton compatibility, motion realism, and avatar fidelity. The presented demo provides a hands-on illustration of how these challenges can be combined into a single real-time system accessible to non-experts. The setup enables participants to see how markerless motion capture can seamlessly animate different digital characters, from stylized avatars to high-fidelity volumetric humans, highlighting practical issues such as skeleton retargeting, rigging, and temporal consistency. In a period where virtual humans play a key role in immersive media, XR, and simulation, this demo shows how low-cost equipment and efficient tracking algorithms can deliver engaging, high-quality results suitable for research, education, and prototyping.

II. DEMO DESCRIPTION

The demo guides attendees through the full real-time pipeline shown in Figure 1. The process begins with markerless body tracking using the ZED SDK: three calibrated ZED 2i cameras capture synchronized RGB-D streams, from which the system extracts 3D joint positions in real time. The multi-view data is then fused to increase robustness, reduce occlusions, and generate a unified body skeleton. Once the pose is estimated, the motion data is streamed directly into an Unreal Engine scene, where users can animate different types of characters. These include the UE Mannequin, Unreal's default modular human model commonly used for game development and prototyping; MetaHuman, a high-fidelity, fully rigged digital human capable of expressive animation; and a real volumetrically captured human avatar reconstructed from a professional multi-camera capture studio. Since volumetric scans are typically unrigged, this character was automatically rigged using AccuRig, making it compatible with the real-time animation pipeline.

A key feature of the demo is active user participation. Attendees are invited to step into the three-camera capture volume and perform movements, which are instantly transferred to the selected digital character. This allows them to experience the challenges and advantages of real-time motion tracking, rigging, and cross-character retargeting.

III. HARDWARE SETUP.

The system runs on a compact but high-performance configuration composed of three ZED 2i stereo cameras arranged around the capture area to minimize occlusions and ensure complementary viewpoints. The processing is handled by a laptop equipped with an NVIDIA RTX 4090 GPU, providing the computational power required for multi-camera depth estimation, real-time body tracking, and live streaming to Unreal Engine. A large external display shows the virtual environment, the animated character, and the reconstructed skeleton, allowing both participants and observers to clearly follow the animation process in real time.