

PersonalEyes: Real-Time Markerless Motion Capture and Gaussian Avatar Generation for Scalable VR Embodiment

Anca Salagean
University of Bath
Bath, United Kingdom
as3101@bath.ac.uk

Samuel Legg
University of Bath
Bath, United Kingdom
sl3168@bath.ac.uk

Crescent Jicol
University of Bath
Bath, United Kingdom
cj406@bath.ac.uk

Mark Townsend
University of Bath
Bath, United Kingdom
mrt64@bath.ac.uk

Eamonn O'Neill
University of Bath
Bath, United Kingdom
maseon@bath.ac.uk

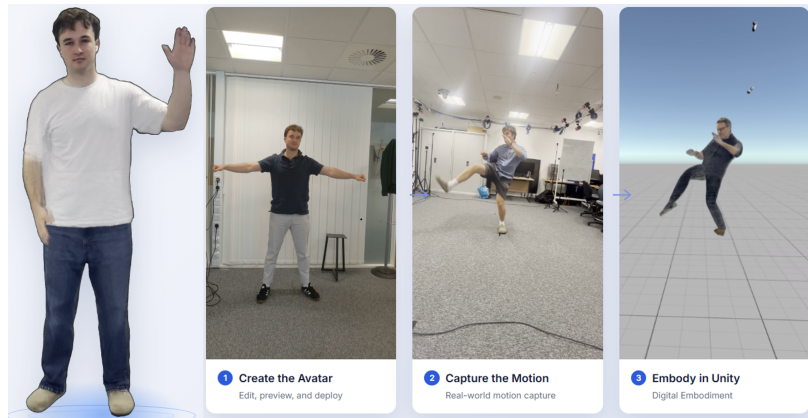


Figure 1: PersonalEyes pipeline: 1) creating a personalised avatar, 2) capturing user motion and 3) real time embodiment in VR.

ABSTRACT

Photorealistic personalized avatars have significant potential for virtual reality (VR), embodiment research and social computing. However, traditional avatar creation workflows require specialist photogrammetry facilities, costly hardware and extensive manual processing. We present PersonalEyes, an end-to-end platform developed at our institution combining IRIS, a real-time markerless webcam-based motion capture system, with a rapid Gaussian avatar generation pipeline running on commodity hardware. Users create a photorealistic animatable avatar from a short monocular video and immediately embody it through markerless motion capture. Attendees will generate, animate and embody their avatars within an immersive VR environment.

AUTHOR KEYWORDS

Gaussian Avatars; Markerless Motion Capture; Virtual Reality.

DOI: 10.37789/icusi.2026.13

Introduction

Avatar appearance influences how users perceive and experience immersive virtual environments [5]. Personalized photorealistic avatars can increase embodiment, presence and emotional and physiological responses [3–6], but their production typically requires expensive photogrammetry, marker based motion capture, specialist expertise and extensive post processing [5].

Recent advances enable photorealistic animatable avatars to be reconstructed from monocular video [2, 7], while markerless motion capture using commodity cameras has reduced barriers to full body tracking [1]. However, these capabilities remain fragmented across separate pipelines, limiting rapid generation and immediate embodiment of personalized photorealistic avatars using low cost hardware.

To address this gap, we present PersonalEyes, an end to end avatar embodiment platform that integrates two in house technologies: a rapid Gaussian avatar generation pipeline and the IRIS real time markerless motion capture framework. Together, these systems enable users to create, animate and embody a personalized photorealistic avatar within a single session using commodity hardware. This paper presents the first public demonstration of the complete PersonalEyes platform.

System Overview

PersonalEyes is an end to end avatar embodiment platform that integrates a Gaussian avatar generation pipeline with the IRIS markerless motion capture framework (Fig. 1). Together, these components enable users to generate, animate and embody a personalized photorealistic avatar within a single session using two to four commodity webcams (OBSBOT MEET SE). Avatar generation takes approximately five minutes from video to Unity-ready avatar on an RTX 4090 workstation, including approximately two minutes of preprocessing followed by three minutes of gaussian optimisation. Once generated, the avatar renders in Unity at up to

70fps. The same camera setup is used throughout the entire workflow for both avatar generation and subsequent real time full body markerless motion capture, eliminating the need for dedicated scanning hardware, motion capture systems or server based processing.

Gaussian Avatar Generation: The avatar generation pipeline reconstructs a personalized photorealistic animatable Gaussian avatar from a short monocular video. The Gaussian avatar representation is based on established deformable Gaussian avatar approaches designed to preserve subject-specific appearance and support realistic novel-pose rendering. Dataset preparation comprises automated person detection, image cropping, portrait matting, pose estimation and body model fitting. These outputs are subsequently used to optimize a deformable Gaussian representation capable of synthesizing photorealistic renderings from novel body poses and camera viewpoints. Following optimization, the reconstructed avatar is exported directly into Unity, where a custom Gaussian rendering pipeline enables real time rendering and interaction. The complete workflow is fully automated, requires only the input from the commodity webcam setup and can be completed within a few minutes. Generated avatars achieve visual quality comparable to commonly used Mixamo ones while preserving user appearance. Quantitative and perceptual evaluation against existing avatar generation methods is ongoing.

IRIS Markerless Motion Capture: IRIS is a multi camera orchestration platform that performs real time markerless whole body motion capture using the same commodity webcam setup employed for avatar generation. Each camera performs GPU accelerated image capture while timestamping every frame. A synchronization layer assembles temporally aligned frame batches, enabling robust multi view processing. During system initialization, camera intrinsics and extrinsics are estimated automatically, removing the need for specialist calibration procedures. Person detection is performed using YOLOX, identities are maintained using OSNet based re identification and whole body pose estimation is generated using RTMW for each synchronized camera stream. The resulting multi view observations are triangulated to reconstruct temporally consistent 3D joint positions, and subsequently filtered and streamed to external applications.

Integrated Embodiment Pipeline: The primary contribution of PersonalEyes is the integration of our in-house Gaussian avatar generation and markerless motion capture systems into a unified embodiment pipeline. Once a personalized photorealistic Gaussian avatar has been generated, IRIS streams real time full body pose estimates directly into Unity, where the reconstructed skeletal motion is applied to the avatar for live animation. By using the same low cost camera setup throughout the entire pipeline, the system enables immediate embodiment of a personalized photorealistic avatar within a single session, without requiring motion capture suits, specialist tracking hardware, dedicated scanning facilities, manual rigging or server based processing. This substantially reduces the cost and complexity of deploying personalized photorealistic avatars, enabling scalable deployment for immersive research and industrial applications.

Demonstration

The demonstration follows a use case in which an attendee creates a personalized photorealistic avatar from a short video

and immediately embodies it in VR within an emotion inducing virtual environment using markerless motion capture (Fig. 1). Attendees will:

- Capture a short video using four standard webcams for Gaussian avatar generation.
- Observe automated preprocessing and reconstruction, including person detection, portrait matting, pose estimation, body model fitting and Gaussian optimization.
- Monitor the optimization process through a live avatar viewer, visualizing the progressive reconstruction of a personalized photorealistic Gaussian avatar.
- Import the generated avatar directly into Unity using the custom Gaussian rendering pipeline.
- Embody and control the personalized photorealistic avatar through a VR head mounted display using live full body markerless motion capture streamed from IRIS. This demonstrates real time avatar animation without motion capture suits or specialist tracking hardware, while spectators can simultaneously observe the immersive experience through a live external display.

Applications

PersonalEyes supports a range of applications in research and industry. By substantially reducing cost and setup complexity, the system broadens access to technologies that have traditionally been restricted to specialist laboratories and commercial studios. Through this demonstration, we showcase how rapid avatar creation and real-time embodiment can be achieved within a single session, opening new opportunities for immersive research, telepresence and digital self-representation.

References

- [1] Saeid Edriss, Cristian Romagnoli, Lucio Caprioli, Vincenzo Bonaiuto, Elvira Padua, and Giuseppe Annino. 2025. Commercial vision sensors and AI-based pose estimation frameworks for markerless motion analysis in sports and exercises: a mini review. *Frontiers in Physiology* 16 (2025), 1649330.
- [2] Zhiyin Qian, Shaofei Wang, Marko Mihajlovic, Andreas Geiger, and Siyu Tang. 2024. 3dgs-avatar: Animatable avatars via deformable 3d gaussian splatting. In *IEEE/CVF*.
- [3] Anca Salagean, Darren Cosker, and Danaë Stanton Fraser. 2025. The Honest Virtual Self? Effects of Avatar Personalization and Motor Control on Physiological Responses to Deceptive Behaviours. In *ISMAR*. IEEE.
- [4] Anca Salagean, Eleanor Crellin, Martin Parsons, Darren Cosker, and Danaë Stanton Fraser. 2023. Meeting your virtual twin: Effects of photorealism and personalization on embodiment, self-identification and perception of self-avatars in virtual reality. In *CHI*.
- [5] Anca Salagean, Daniel A Ledger, Dominic Potts, Tarini Sehgal, Sarah Binesmael, Koner Kalkanel, Hussain Karim, Michelle Wu, George Fletcher, Eamonn O'Neill, et al. 2026. Virtual Selves, Real Emotions: The Impact of Photorealistic Avatar Personalization on Emotion in Virtual Reality. In *CHI*.
- [6] Anca Salagean, Michelle Wu, George Fletcher, Darren Cosker, and Danaë Stanton Fraser. 2024. The utilitarian virtual self—using embodied personalized avatars to investigate moral decision-making in semi-autonomous vehicle dilemmas. *TVCG* (2024).
- [7] Zhijing Shao, Zhaolong Wang, Zhuang Li, Duotun Wang, Xiangru Lin, Yu Zhang, Mingming Fan, and Zeyu Wang. 2024. Splattingavatar: Realistic real-time human avatars with mesh-embedded gaussian splatting. In *IEEE/CVF*.