Volume 16 • Issue 3 • 2024

Engineering Management in Production and Services

sciendo

ISMSME International Society for Manufacturing, Service and Management Engineering

received: 16 February 2024
accepted: 15 August 2024

pages: 1-14

Digital twin-based alternate ego modeling and simulation: Eva Herzigová as a 3D MetaHuman avatar

George Lăzăroiu
Tom Gedeon
Danuta Szpilko
Katarzyna Halicka

ABSTRACT
Due to Dimension Studios and Unsigned Group, by likeness and movement digitization, the human supermodel Eva Herzigová's highly realistic metahuman and 3D digital twin lead to perpetual career extension (e.g., on virtual runway shows by catwalk replication or in advertising campaigns by particular pose adoption), including in virtual worlds such as metaverse spaces, by use of extended reality technologies, in terms of virtual clothing, digital hair, skin texture, facial and body expressions, and makeup styling as avatar appearance and persona. In this paper, we show how, by likeness capturing (e.g., signature walk by motion capture shoot for face and body movement tracking) and virtual clothing, Eva Herzigová's hyperreal 3D avatar (lifelike 3D digital human clone or realistic virtual human) is configured with subsequent restyled hair and make-up, 3D clothing modeling, and extended reality fashion shows on virtual catwalks or campaign shoots. We clarify that photographic reference materials and a sensor-based motion capture suit assisted Metahuman Creator in capturing facial and motion data for Eva Herzigová's digital avatar, configuring realistic depiction and virtual activations of beauty standards, in addition to digital likeness, signature style, talent representation, and narrations, typifying digital human creation and presence in virtual worlds (e.g., for metaverse-based shoots). We conclude that Eva Herzigová's hyperrealistic 3D digital human twin and MetaHuman avatar operate as a styled virtual alter ego and versatile personal brand extension by motion capture technologies for immersive virtual fashion shows, live video streams and broadcast, and digital wearable modeling.

KEY WORDS
Eva Herzigová, digital twin, hyperrealistic 3D digital human twin, MetaHuman avatar appearance and persona, motion capture technologies, artificial intelligence, AI

10.2478/emj-2024-0020

George Lăzăroiu
Curtin University, Kent Street, Bentley Western Australia 6102, Australia
Toronto Metropolitan University, 350 Victoria Str., ON M5B 2K3 Toronto, Canada
Cardiff Metropolitan University, Llandaff Campus, Western Avenue, CF5 2YB Cardiff, United Kingdom
Spiru Haret University, 13 Ion Ghica Str., District 3, Bucharest, Romania
ORCID 0000-0002-3422-6310

Corresponding author:
e-mail: phd_lazaroiu@yahoo.com

Tom Gedeon
Curtin University, Kent Street, Bentley Western Australia 6102, Australia
ORCID 0000-0001-8356-4909
e-mail: tom.gedeon@curtin.edu.au

Danuta Szpilko
Bialystok University of Technology, Wiejska 45A, 15-351 Bialystok, Poland
ORCID 0000-0002-2866-8059
e-mail: d.szpilko@pb.edu.pl

Katarzyna Halicka
Bialystok University of Technology, Wiejska 45A, 15-351 Bialystok, Poland
ORCID 0000-0001-6370-8995
e-mail: k.halicka@pb.edu.pl

Introduction

3D image generation and motion segmentation algorithms, image-based 3D reconstruction and data acquisition technologies, and physio- and neuro-behavioral sensors can be deployed in 3D facial expression synthesis, perceptual and cognitive performance, expressive body language and movements, emotional and personality features, and interactive social behaviors. Sensor network clustering and data augmentation algorithms, 3D clothing simulation and virtual try-on technologies, and Internet of

Lăzăroiu, G., Gedeon, T., Szpilko, D., & Halicka, K. (2024). Digital twin-based alternate ego modeling and simulation: Eva Herzigová as a 3D MetaHuman avatar. *Engineering Management in Production and Services, 16*(3), 1-14. doi: 10.2478/emj-2024-0020

Things (IoT)-based machine vision and image sensor devices can be leveraged in multi-sensory object localization, recognition, and representation, realistic 3D avatar facial appearance animation, and photo-realistic synthetic image and scene-aware 3D human avatar pose generation. Mediated virtual self-expression and representation, avatar-mediated self-disclosures and social identification, and facial movement detection and recognition develop on multi-modal image-based 3D object detection and scene perception algorithms, IoT-connected wearable and spatial computing devices, and human-like tactile sensing and perception technologies.

Due to Dimension Studios and Unsigned Group, by likeness and movement digitization, the human supermodel Eva Herzigová's highly realistic metahuman and 3D digital twin lead to perpetual career extension (e.g., on virtual runway shows by catwalk replication or in advertising campaigns by particular pose adoption), including in virtual worlds such as metaverse spaces, by use of extended reality technologies, in terms of virtual clothing, digital hair, skin texture, facial and body expressions, and makeup styling as avatar appearance and persona.

Real-time model volumetric scanning, augmented reality try-on technologies and filters, and 3D real-time design can be harnessed in performance capturing and hologram streaming for alternate egos and personal brand extension as digital personalities and virtual counterparts by accurate lifelike features and authentic expression, without body and face-enhancing alterations (e.g., skin smoothing, eye widening, lip plumping, or lash darkening).

1. Literature review

3D spatial computing and metaverse technologies, facial shape and appearance modeling, and spatial mapping algorithms (Al-Emran & Deveci, 2024; Grupac & Lăzăroiu, 2022) enhance lifestyle behavior actions and habits of 3D photorealistic avatars across engaging extended reality environments. 3D garment simulation technologies improve multi-dimensional metaverse experiences, behaviors, and preferences throughout avatar motion and virtual scene control, image-based 3D scene representation, and 3D clothed avatar and facial appearance reconstruction.

Animatable 3D virtual human avatar tools (Bugaj et al., 2023; Zhang et al., 2023a) shape 3D-aware virtual clothed human generation, reconstruction, and animation, image-based autonomous virtual 3D object recognition, segmentation, classification, and modeling, and generative 3D-aware human face and image synthesis. 3D-aware human facial appearance and expression analysis, 3D scene representation technologies, and animatable neural human avatar generative modeling optimize controllable 3D-aware human avatar generation and image synthesis, reconstruction, and animation, single image-3D pose mapping and estimation, and 3D human representation decomposition. Animatable 3D-aware clothed human generation and modeling enhance photo-realistic human image, pose, and shape synthesis, virtual body appearance variations, and 3D neural object and scene representation.

Deep learning visual and speech recognition algorithms (Zhang et al., 2023b; Zvarikova et al., 2023) configure contextual deep convolutional neural network-based facial feature and body shape virtualization, single image-based 3D human clothed avatar reconstruction and pose estimation, and 3D human body texture, shape, and pose estimation, tracking, and prediction. Visual image and geospatial mapping algorithms articulate skin texture reconstruction, simulation data reconstructed appearance, and 3D garment texture visualization. Perceptual image algorithms and 3D avatar motion capture technologies assist in 3D human pose and physics-based appearance estimation, garment and hairstyle simulation, and human body avatar modeling.

Computer graphics and affective haptics technologies, multimodal visual interaction and voice recognition tools, and computer animation and speech synthesis techniques (Cui & Liu, 2023; Valaskova et al., 2022) assist in motion capture animation, personalized voice expression and visual appeal, and realistic texturing and lighting. Natural speech recognition and voice generation technologies enable digital human image construction in terms of hyper-realistic anthropomorphic appearance, facial expression and emotion recognition, and real-time interactive feedback capabilities. Image and voice recognition technologies, synthetic data augmentation, and social simulation emotion modeling further realistic facial image reconstruction, physiological and behavior data sharing and interoperability, and object visual and body shape appearance. Hyper-realistic immersive experiences can be attained by use of motion pose capture and automatic speech recognition technologies, image-based temporally coherent visual and spatial semantic representation, and visual perception and semantic object recognition algorithms.

Physics-based clothing animation and simulation, full-body motion capture analysis, and animatable full-body appearance modeling (Kliestik & Lăzăroiu, 2023; Xiang et al., 2022) configure pose-driven full-body avatars, neural rendering-based physically-simulated garments and photorealistic clothing, and full-body avatar representation and movement predictions. Photorealistic neural clothing appearance modeling and 3D human body pose and shape reconstruction and estimation articulate emotionally expressive animatable photorealistic full-body avatars, elaborate body–clothing interactions, and virtual garment resizing. 3D human avatar and realistic clothing modeling and physically-realistic clothing simulation assist in photorealistic avatar garments and clothing appearance, avatar self-expression, and photorealistic synthetic human representation and clothing animation.

2. Methodology

Realistic avatar and remote sensing technologies, deep learning-based visual reasoning and sentiment analysis, and behavioral pattern and sensor data context-aware interpretation enable 3D face reconstruction and modeling, deep learning-based semantic segmentation, and body shape and emotional avatar perception across 3D virtual simulation environments. Speech visualization and tactile sensor technologies articulate 3D garment and visual object appearance modeling, complex adaptive IoT network dynamical behavior simulation, and multisensor data fusion and integration. Visual recognition and multilayer perceptron algorithms assist in head, hand, eyebrow, eye, and lip movement tracking and synchronized capture, visual motion awareness, and 3D spatial object modeling across visually-consistent photo-realistic virtual environments.

Computer vision and graphics techniques, social interaction and environment data, and sensing and perception systems further dynamic deep generative and neural scene modeling, visual data processing, and semantic scene understanding. Semantic reasoning and multi-view imaging techniques, lifestyle behavior and body shape modeling, and deep fusion convolutional neural networks shape animatable 3D clothed human and 4D dynamic scene reconstruction across 3D spatial data analytics and infrastructures. Semantic ontology and multimodal sensing technologies, visual language navigation and semiotics, and human digital twin body and behavior modeling optimize image visual coherence, mood and emotion recognition, and body shapes and sizes. Space mapping and visual perception algorithms, visual object appearance modeling, and 3D facial expression and motion tracking enhance physical condition detection and monitoring and remote sensing image scene classification, processing, and fusion by use of synthetic motion and perception data. 3D human body and image appearance modeling improves remote sensing image classification, physical world attributes and states, and multiscale knowledge mining.

Photographic reference materials and a sensor-based motion capture suit assisted Metahuman Creator in capturing facial and motion data for Eva Herzigová's digital avatar, configuring realistic depiction and virtual activations of beauty standards, in addition to digital likeness, signature style, talent representation, and narrations, typifying digital human creation and presence in virtual worlds (e.g., for metaverse-based shoots). Our analysis covers reality capture, mocap shoot, Z Brush, clothing simulation, hair creation, make up, skin and clothing texture, unreal animation, environment build, and lighting pass.

3. Results

3D cloth modeling and simulation (Valaskova et al., 2022; Xiang et al., 2022) enable multi-view clothing capture, photorealistic texture-based clothed body shape estimation, and physically-simulated 3D contextual detail-aware interactive virtual clothing animation by use of reality capture technologies. 3D virtual fabric simulation, silhouette image analysis, and structure-preserving 3D and context-aware virtual garment modeling further physically-inspired clothing appearance, photorealistic clothing texture, and 3D human skeleton motion prediction, recognition, capture, generation, analysis, tracking, segmentation, and representation with regard to photorealistic animatable clothed avatars. Photorealistic clothing and body motion animation, personalized photorealistic avatar generation and representation, and photorealistic texture and clothing appearance generation can be attained by generative artificial intelligence (AI) synthetic and kinematic

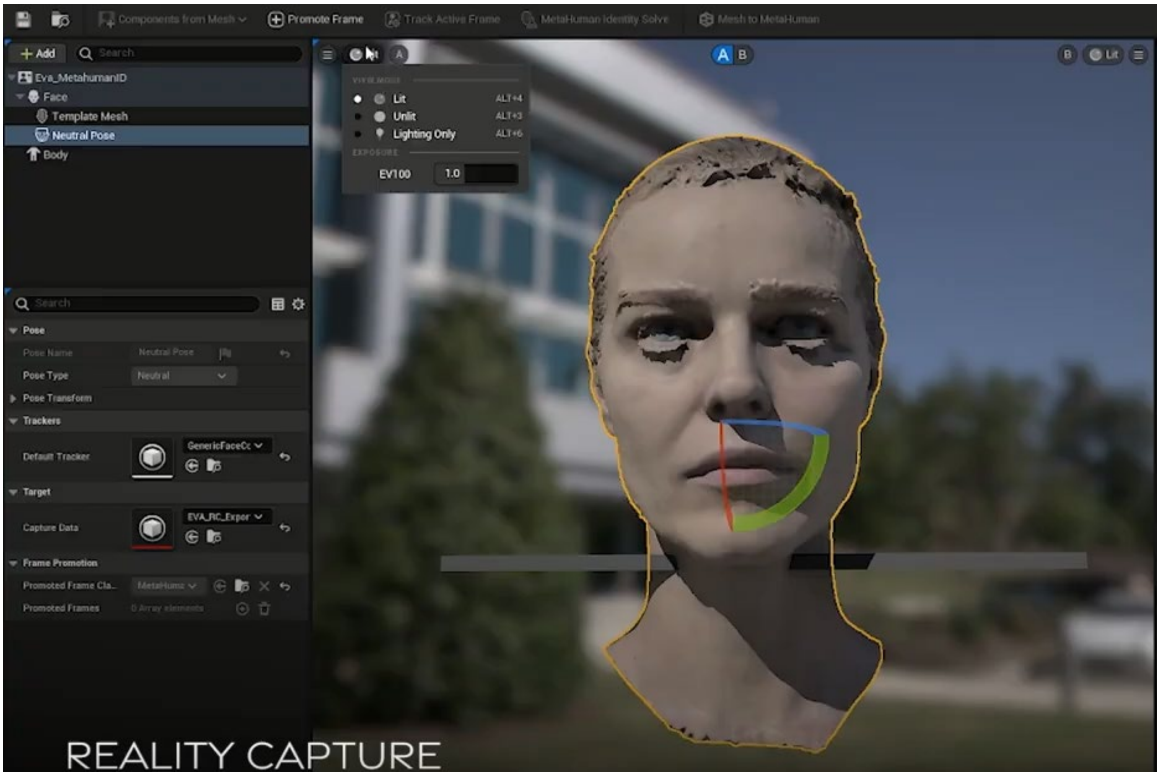

Fig. 1. Reality capture
Source: © Dimension Studio and Unsigned Group.

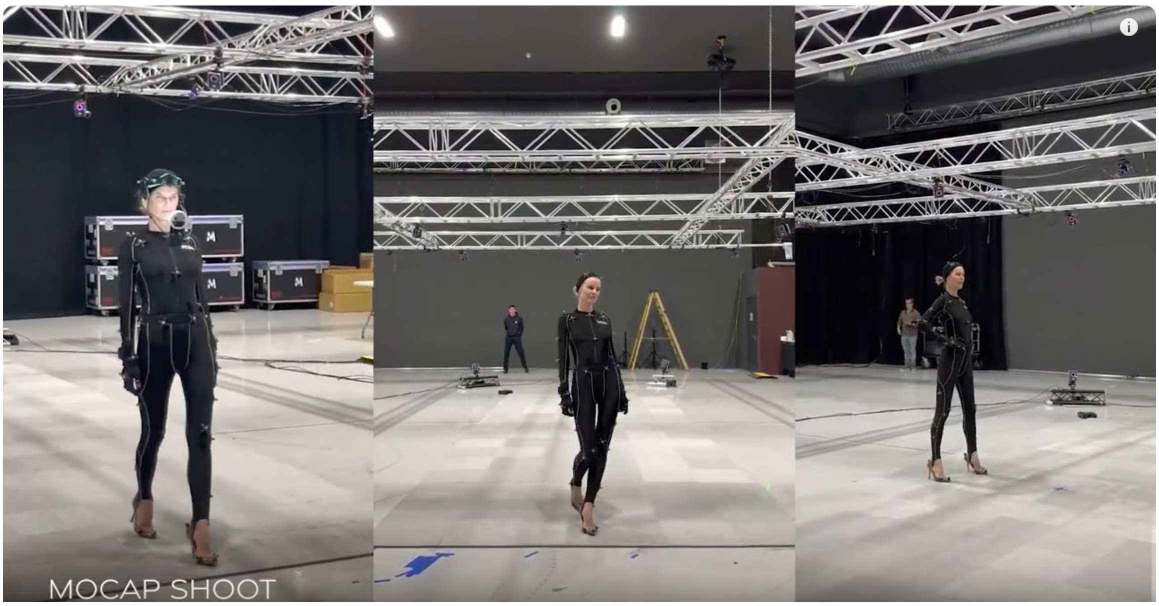

Fig. 2. Mocap shoot
Source: © Dimension Studio and Unsigned Group.

body pose data, dynamic gesture, multi-modal emotion, and skeleton-based action recognition algorithms, and 3D garment and photorealistic appearance modeling.

Neural 3D clothed-human body reconstruction, 3D realistic skin and hair modeling, and autonomous simulated virtual agent intentionality (Bugaj et al., 2023; Wang et al., 2024) can be achieved by 3D motion data processing and analysis, sensor-based 3D semantic mapping, and syntactic navigation and facial expression data with regard to digital human appearance and photo-realistic 3D clothed virtual humans. Cognitive mapping algorithms, real-time subtle facial expression capture, and 3D human shape representation and modeling shape photo-realistic 3D human and synthetic clothing reconstruction in 3D immersive visualization environments. Digital twin mapping and immersive technologies, synthetic human data, and 3D motion capture systems optimize 3D facial expression and skin texture modeling and 3D machine learning-based simulation in immersive virtual reality simulation environments (Figure 1).

Visual sensor and motion detection data, 3D digital twin visualization technologies, and activity-tracking algorithms (Grupac & Lăzăroiu, 2022; Jiang et al., 2024) improve intelligent wearable garment, deep learning 3D limb movement and vision sensor-based human tracking, and human pose estimation with regard to mocap shoots. Image sensor and gesture recognition technologies, deep learning sensor fusion and neuromorphic computing algorithms, and intelligent 3D garment and emotion recognition systems can be harnessed in 3D skeletal rendering, deep learning network-based facial expression recognition, and 3D hand and leg tracking. Visual imaging technologies, garment sensor networks, physiological and motion data, and wearable textile sensors can be deployed in wearable sensor clothing, 3D sensor data mapping and virtual scene generation, spatial movement patterns, human body movement reconstruction, and facial feature detection. Machine learning and computer vision algorithms and space syntax simulation (Chen et al., 2021; Valaskova et al., 2022) enable convolutional neural network-based facial expression, hand gesture, and body shape, and action recognition modeling, photorealistic deep neural network and sensory data-based 3D digital human reconstruction, and image-based 3D object reconstruction and scene representation. 3D visualization and full body motion capture systems and 3D generative AI simulation technologies further generative AI digital twin-based virtual reality simulation and socio-spatial modeling by sensor-based physiological data and unstructured image collection, image representation reconstruction, and 3D semantic occupancy and navigational pattern predictions in relation to mocap shoots. Networked social-emotional interactivity, sensor-based situational awareness, deep learning-based 3D human representation, complex immersive spatial and social dynamics simulation, and photorealistic image generation and texture rendering can be attained by virtual 3D and computer-based visual simulation technologies (Figure 2).

Human–machine interaction and sensor network technologies (Fjørtoft Ystgaard et al., 2023; Zvarikova et al., 2023) can be leveraged in immersive augmented reality visualization, social emotion, behavior, intention, and attitude mapping, and contextual environment understanding in relation to human digital twins across human-centric IoT network infrastructures by use of Z Brush. Embodied interaction modalities, interconnected quantified selves, design process participatory involvement, collaborative task performance assessment, algorithmic social behavior, and meaningful human-like features develop on sensors and actuators, algorithmic modeling and semantic mapping techniques, shared knowledge and action mechanisms, and face recognition and wireless sensor network technologies with regard to artificial social peers across pervasive IoT environments. Immersive extended reality technologies, photorealistic avatar modeling, and perceptual image processing and spatial cognition algorithms (Alldieck et al., 2022; Grupac & Lăzăroiu, 2022) enhance deep neural network-based photorealistic 3D human visualization and reconstruction by object–scene semantics, immersive virtual apparel try-on, 3D virtual clothing, and realistic 3D avatar animation. Photorealistic digital human modeling and immersive 3D human body shape and pose reconstruction improve emotionally expressive dressed 3D avatars in terms of predicted appearance and skin tone perceptual quality, virtual garment rendering, and visual motion perception by use of Z Brush. Deep learning-based semantic segmentation, 3D scene capturing and object reasoning techniques, 3D facial appearance modeling, and multiscale shape representation and analysis can be harnessed in 3D scene and clothed human reconstruction and human-like avatar representation (Figure 3).

Physical appearance, facial expression, body shape and movements, voice recognition, and social personality, together with emotional and physiological state, body movement, and natural speech replication, with regard to multiple fusion technology-based virtual human individuals (Cui & Liu, 2023; Kliestik et al., 2022), integrate 3D deep learning data-driven reconstruction and scanning techniques. Human appearance and communication, emotion recognition and expression, facial and body movement, and interaction behavior simulation require emotion capture and recognition technologies, convolutional neural network and visual perception algorithms, sensor-based gesture capture and eye-tracking devices, emotional speech synthesis systems, and environmental behavior monitoring sensors. Natural language processing and speech synthesis necessitate

Fig. 3. Z Brush
Source: © Dimension Studio and Unsigned Group.

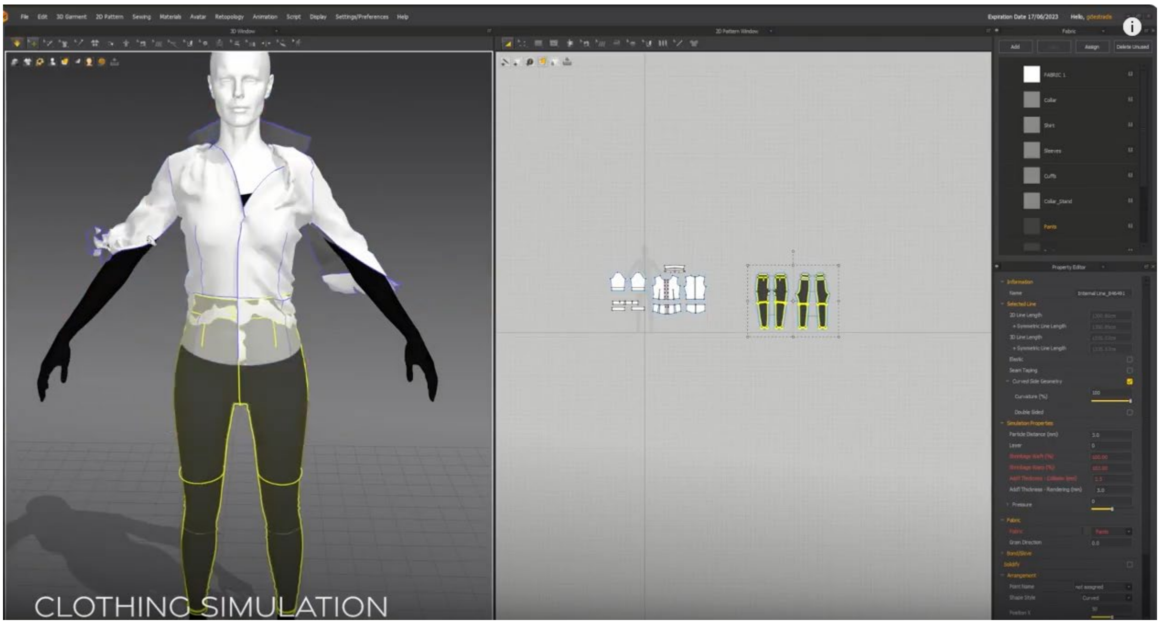

Fig. 4. Clothing simulation
Source: © Dimension Studio and Unsigned Group.

deep learning-based face image mapping and digital human appearance and emotional expression simulation for 3D modeling technology-based digital humans. Computerized facial animation and motion control technologies, deep and machine learning algorithms, and image-based 3D reconstructed scene segmentation are pivotal in multimodal human interaction behavior simulation and modeling in immersive augmented reality and anthropomorphic simulation environments. Convolutional neural network deep learning 3D pose estimation and body shape visualization technologies (Correia & Brito, 2023; Zvarikova et al., 2023) can be deployed in computer vision-based 3D human body reconstruction, 3D human silhouette and skeleton modeling, and realistic 3D body and object representation. 3D human body reconstruction technologies, deep learning prediction and 3D location sensing algorithms, and spatial location mapping can be leveraged in image-based 3D clothed human body shape reconstruction, multi-person 3D motion prediction and pose estimation, and multi-view convolutional neural network-based distributed sensing, image data classification, and 3D object retrieval and shape recognition. Deep learning-based expressive face and clothing texture recognition, 3D deep neural network-based facial expression modeling, and 3D human body shape and pose appearance, estimation, tracking, and prediction develop on digital human modeling algorithms, silhouette image recognition and analysis, and 3D body segmentation. Convolutional neural network-based spatial scene recognition, multi-view 3D scene and object reconstruction, and 3D garment mapping integrate photorealistic image stylization and multi-object 3D reconstruction algorithms. Virtual anthropomorphic avatars require multi-view 3D human body shape and pose reconstruction, facial expression control, and synthetic human data (Figure 4).

Extended reality and metaverse technologies, digital twin-based mobile edge computing and visual attention networks, and deep reinforcement learning algorithms (Yu et al., 2024; Zvarikova et al., 2023) can be harnessed in deep learning-based 3D face and photorealistic skin appearance reconstruction by use of perceived humanness and environmental data in relation to digital hair creation. 3D virtual try-on and garment simulation technologies, visual motion perception systems, and semantic human body tracking can be deployed in 3D human pose estimation and neural cloth simulation, machine learning-dynamic scene 3D reconstruction, and 3D body motion capture and realistic garment generation, optimizing immersive interactive experiences. Avatar skin tone, 3D garment and physics-based clothing animation, garment texture image segmentation and design classification, and virtual avatar appearance (Kliestik et al., 2022; Sarakatsanos et al., 2024) necessitate 3D avatar garment and fashion design simulation, cloth and body simulation techniques, and virtual try-on clothing systems with regard to digital hair creation. Motion tracking algorithms are pivotal in 3D virtual clothing and digital garment design visualization, anthropomorphic avatar facial expression recognition, avatar perception processes, and avatar motion capture and control in immersive extended reality environments. Physics-based 3D cloth and garment behavior simulation and image-based 3D object modeling are instrumental in 3D digital garment design, immersive virtual avatar garment fitting, anthropomorphic avatar body movements and gestures, and avatar emotion detection and recognition (Figure 5).

3D virtual garment simulation technologies, 3D garment modeling and body motion prediction algorithms and photorealistic appearance and physics-based dynamic modeling (Bertiche et al., 2022; Kliestik et al., 2022) can be leveraged in 3D garment animation generation and physics-based differentiable rendering with regard to digital make-up. 3D human body pose detection, estimation, recognition, and tracking develop on physics-based appearance and context-aware garment modeling, 3D neural garment animation simulation technologies, and deep learning-based neural cloth simulation. 3D virtual clothing simulation and garment appearance integrates 3D cloth animation and body motion augmentation techniques, 3D avatar modeling, and deep 3D human pose estimation and object detection networks. 3D immersive realistic sensory experiences (Bugaj et al., 2023; Cui & Liu, 2023) can be achieved by use of virtual human and image perception technologies, 3D object motion prediction, estimation, perception, detection, trajectory, and tracking, and deep neural network and machine intelligence algorithms. 3D modeling and computer virtual simulation technologies, visual perception and semantic data augmentation algorithms, and photorealistic 3D object and scene generation shape facial emotional expression of realistic-like virtual humans by automated meaning extraction and 3D image reconstruction in relation to digital make-up. Virtual human and brain-computer interface technologies, deep learning-based cognitive simulation and image segmentation algorithms, and robotic simulation and cognitive sharing systems optimize avatar-mediated communication in terms of personality attribution and social identification for facial expression and body movements (Figure 6).

3D personalized body shape and face appearance customization, 3D clothed human avatar and skin surface reconstruction, and 3D facial pose estimation, variation, and tracking (Ding et al., 2024; Kliestik & Lăzăroiu, 2023) require deep neural network-based facial animation, body shape, skin and hair color, and human pose modeling, facial texture recognition and analysis algorithms, and computational visual attention and multimodal generative AI systems with regard to skin and clothing texture. Facial and body texture recognition for personalized 3D human avatars necessitate generative adversarial network-based animatable 3D-aware, multi-view frontal, and controllable face image production, deep learning-based gaze and expression estimation algorithms, and image-based 3D facial shape reconstruction. Face angle detection and image data analysis algorithms, image acquisition technologies, and 3D facial expression modeling are pivotal in photo-real-

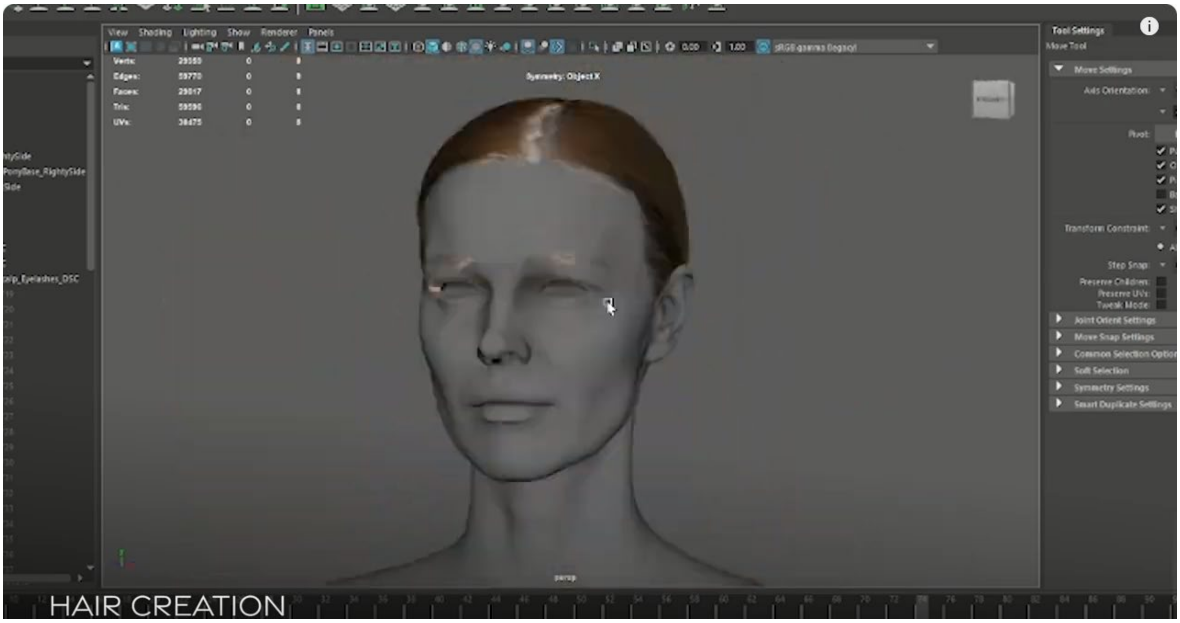

Fig. 5. Hair creation
Source: © Dimension Studio and Unsigned Group.

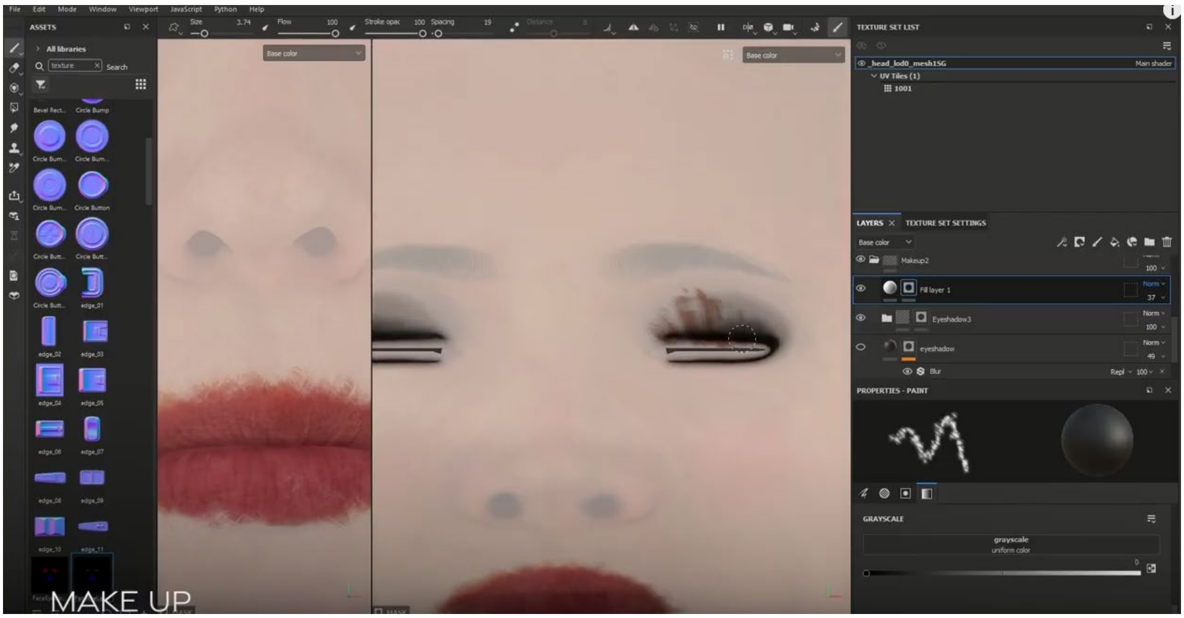

Fig. 6. Make up
Source: © Dimension Studio and Unsigned Group.

istic facial texture generation, analysis, reconstruction, and transfer. 3D head and facial shape prediction, 3D face and body shape modeling, facial texture and expression reconstruction, and computational visual and perceptual algorithms are instrumental in skin texture appearance and facial appearance and texture. Human behavior-recognition and visual perception algorithms (Bugaj et al., 2023; Lee & Joo, 2023) enable human-object interaction detection and recognition, neural radiance field scene and object-centric 3D representation, and photorealistic synthetic image generation and 3D scene reconstruction in relation to skin and clothing texture. 3D path planning and visual motion algorithms assist in human motion capture data segmentation, processing, and analysis, image-based 3D object and scene generation, representation, reconstruction, and modeling, and 3D virtual body formation, size, image, and shape perception (Figure 7).

Neural radiance field- and physics-informed deep learning-based skeletal motion recognition, capture, and prediction, computer vision-based 3D

Fig. 7. Skin and clothing texture
Source: © Dimension Studio and Unsigned Group.

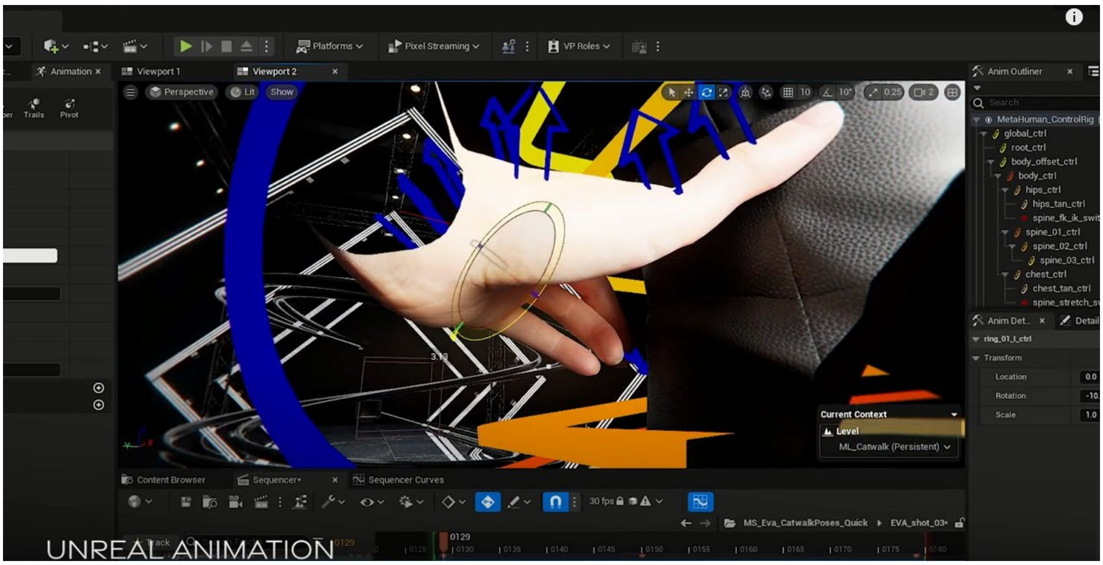

Fig. 8. Unreal animation
Source: © Dimension Studio and Unsigned Group.

semantic scene segmentation, modeling, and mapping, and facial animation and clothed human reconstruction algorithms (Liao et al., 2023; Valaskova et al., 2022) enable virtual body appearance throughout animatable human avatar modeling and reconstruction by use of Unreal animation sequences. 3D facial and image appearance modeling and controllable topology-consistent multi-view imagery further virtual facial expression recognition, photorealistic clothed virtual human images, and computer vision-based 3D object detection, localization, detection, positioning, and tracking. Avatar immersive experiences can be attained by use of deep learning-based 3D skeleton motion prediction, recognition, capture, retargeting, tracking, and segmentation, clothed 3D human shape reconstruction and avatar modeling, and single image-based 3D clothed avatar creation and reconstruction. Neural 3D object detection, representation, segmentation, reconstruction, modeling, and tracking, deep learning-based 3D object detection and semantic segmentation virtual multi-view fusion, and 3D human body shape prediction and modeling (Kliestik & Lăzăroiu, 2023; Liao & Waslander, 2024) require 3D face recognition and multiple object tracking algorithms. Deep learning semantic scene understanding, skeleton motion pre-

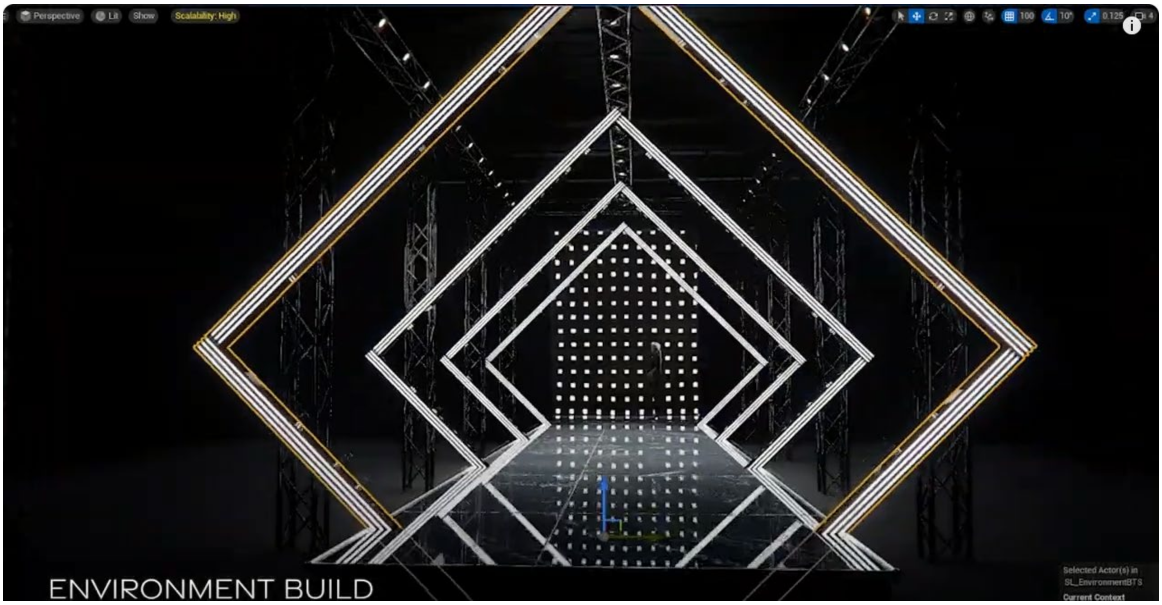

Fig. 9. Environment build
Source: © Dimension Studio and Unsigned Group.

Fig. 10. Lighting pass
Source: © Dimension Studio and Unsigned Group.

diction, recognition, localization, capture, retargeting, segmentation, representation, and tracking, and 3D object and scene representation necessitate 3D human pose estimation and image semantic segmentation algorithms, together with Unreal animation sequences. 3D visual computing and holographic telepresence technologies are pivotal in multi-view fusion neural network and object detection, 3D scene decomposition and segmentation, neural 3D scene and object shape reconstruction, deep learning 3D object shape prediction, 3D image decomposition, and neural object representation (Figure 8).

Semantic image modeling techniques, 3D-aware generative adversarial networks, and controllable clothed animatable human avatar generation (Grupac & Lăzăroiu, 2022; Zhang et al., 2023a) improve identity and appearance synthesis across the environment build. Animatable 3D-aware clothed human generative modeling, 3D image data visualization and remote sensing technologies, and 3D visual object recognition and semantic perception algorithms can be harnessed in facial pose synthesis, 3D scene visualization and representation, and controllable and animatable 3D-aware clothed human avatar generation. Social interaction and human motion modeling techniques can be deployed in 3D dynamic visual scene generative and human avatar modeling, human facial appearance simulation, and realistic scene and object representation. Sensor signal interpretation algorithms, multi-sensor data augmentation techniques, and deep 3D neural scene representation and object detection networks (Uhlenberg et al., 2023; Zvarikova et al., 2023) can be harnessed in human digital twin simulation and modeling, perceptually-guided memory-based human motion simulation, and shared cognitive behavior and knowledge learning processes. Gait event detection and eye gaze estimation algorithms, 3D human body motion capture technologies, and gait event estimation and motion capture data analysis optimize computer vision-based body pose augmentation, real-time avatar motion capture, analysis, and simulation, and 3D human silhouette modeling by use of virtual sensor data. Hand and head gesture recognition algorithms, human motion simulation and holographic communication technologies, and anthropometric data capture tools enhance virtual body sensor data processing throughout the environment build. Multi-sensor signal processing and eye-movement event detection algorithms, human digital twin technologies, and sensor-based movement assessment and analysis improve visuo-spatial body shape representation (Figure 9).

3D fully clothed human body reconstruction, multisensor hand tracking, 3D facial expression and scene modeling, and multisensor gesture recognition (Kliestik et al., 2022; Lee et al., 2024) develop on computer graphics and vision technologies in relation to the lighting pass. Spatial feature extraction algorithms can be leveraged in deep learning-based realistic digital 3D avatar creation and garment reconstruction, deep graph convolutional network-based multi-layer clothing modeling, and 3D multi-view facial capture and expression recognition. Human skeletal pose estimation and appearance modeling, remote sensing scene classification, facial expression mapping, and multi-view facial and hairstyle expression integrate 3D motion capture systems. Deep semantic segmentation and image generation algorithms, visual object recognition and understanding systems, and semantic, image, and motion capture techniques (Grupac & Lăzăroiu, 2022; Yang et al., 2023) further animatable 3D avatar generation and representation, realistic 3D human motion synthesis, and animatable human representation reconstruction. Deep semantic segmentation networks, realistic human avatar modeling, and remote sensing imagery semantic segmentation shape facial and body pose estimation, detection, emotion recognition, tracking, reconstruction, and analysis, realistic 3D human motion generation, representation, and prediction, 3D human avatar and neural scene representation, and visual scene understanding with regard to the lighting pass (Figure 10).

4. Discussion

3D convolutional neural network-based dynamic hand gesture and human action recognition, object understanding, multimedia event detection, and image semantic segmentation (Bugaj et al., 2023; Xiang et al., 2022) can be achieved by deep learning physics-based cloth simulation and image semantic segmentation algorithms, clothing shape capture and reconstruction, and physics-based deep multi-layered clothing appearance simulation. Realistic clothing and skin texture modeling, 3D clothing simulation and body avatar generation technologies, and 3D garment and synthetic clothing animation shape photorealistic clothing texture and full-body clothed avatar appearance. Physics-based cloth simulation optimizes data-driven photorealistic clothing appearance modeling, clothed human reconstruction, photorealistic clothing texture generation, and human body behavior recognition and mapping. Clothing data simulation, multi-modal sensor data fusion and network clustering, and object and scene semantics enhance physically-simulated 3D scene and garment in avatar mediated-virtual immersive environments.

3D animation simulation and computer vision-based motion capture technologies, 3D human pose capture and motion tracking algorithms, and speech and emotion recognition analysis (Cui & Liu, 2023; Kliestik & Lăzăroiu, 2023) are instrumental in psychological state and emotion expression simulation, image and voice modeling, and realistic facial expression replication. Computer graphics and virtual human technologies, deep learning facial expression migration algorithms, 3D realistic virtual human simulation, and 4D human body motion and image capture modeling configure AI virtual human interactivity, deep learning 3D face reconstruction, and character motion capture. 3D scanning and modeling techniques, face expression and movement simulation, deep learning 3D character animation generation, and 3D facial image mapping and reconstruction technologies articulate 3D body shape and personality prediction accuracy and immersive interactive virtual human behavior.

Data-driven mobile interaction-aware human motion prediction algorithms are instrumental in deep reinforcement learning-based human-aware 3D scene generation and reconstruction, deep recurrent neural network-based context-aware 3D human motion prediction, and deep learning-based human behavior prediction and recognition. Human-scene and object interaction-aware prediction algorithms (Lee & Joo, 2023; Valaskova et al., 2022) configure 3D scene interaction and understanding, 3D human-scene interaction generation, simulation, and modeling, and 3D clothed human avatar and scene-aware motion generation by use of skeleton-based motion capture data. Human body posture detection and recognition algorithms articulate computer vision-based 3D human motion capture, 3D garment modeling and stylization, 3D textured clothing, 3D human pose forecasting, and realistic 3D human motions.

Machine learning and computer vision algorithms, virtual and augmented reality technologies, and generative AI physics-based modeling (Kliestik & Lăzăroiu, 2023; Lattas et al., 2023) configure hyper-realistic immersive navigational decision-making visualization in shared immersive visual physically-simulated 3D environments. Networked cognitive robotic and computer vision systems, 3D facial and multi-image reconstruction technologies, and computer vision and haptic perceptual algorithms articulate immersive 3D experiences and cognitive modeling and simulation with regard to facial appearance, expression, and shape capture. Physics-informed neural and multi-layered sensor networks, facial shape and knowledge-driven physical modeling, and geolocated event modeling and simulation assist photorealistic facial and texture rendering and skin tone augmentation in 3D immersive virtual environments.

Conclusions

Deep facial emotion expression recognition, visual object shape and appearance reconstruction, and 3D human body shape and texture mapping reconstruction and modeling integrate reinforcement learning-based generative physical AI simulation workflows, neurophysiological, behavioral, and environmental sensors, and edge computing and extended reality technologies. Perceptual and cognitive processes, 3D object localization, and scene recognition and coherence require cloud and edge computing technologies, deep learning-based cognitive digital twins, social interaction behavior simulation and modeling, and geospatial intelligence and human body motion analysis in physically-based virtual worlds. Deep learning-based cognitive state prediction and task recognition, 3D photorealistic avatar representation, and bodily gestures and movements necessitate deep multi-agent reinforcement learning modeling, human skeletal pose and real-time motion trajectory tracking, and real-time body data and human location monitoring.

By likeness capturing (e.g., signature walk by motion capture shoot for face and body movement tracking) and virtual clothing, Eva Herzigová's hyperreal 3D avatar (lifelike 3D digital human clone or realistic virtual human) is configured with subsequent restyled hair and make-up, 3D clothing modeling, and extended reality fashion shows on virtual catwalks or campaign shoots. Eva Herzigová's hyperrealistic 3D digital human twin and MetaHuman avatar operate as a styled virtual alter ego and versatile personal brand extension by motion capture technologies for immersive virtual fashion shows, live video streams and broadcast, and digital wearables modeling.

Limitations and further directions of research

Limitations are associated with covering a single 3D digital human twin and a specific cooperation between two companies creating a MetaHuman avatar. Further research should inspect how moving object trajectory clustering and multilayer perceptron algorithms, multi-modal 3D object detection, image recognition and completion, remote sensing data, and humanoid robotic and wearable sensor technologies are pivotal in human skeleton motion tracking and pose estimation, physics-based clothing and human body behavior simulation, and 3D avatar reconstruction. Subsequent analyses can show how deep learning perceptual and visual object tracking algorithms, context-aware semantic reasoning and image sensing systems, and multi-source body sensor data analysis are instrumental in deep facial emotion expression recognition, physiological and psychological states, and 3D scene understanding and object classification. Thus, it would be of interest to clarify how blockchain-based metaverse and digital twin technologies, distributional semantic and deep reinforcement learning algorithms, human activity and decision-making knowledge modeling, and multiple data source integration configure 3D visual scene representation, reconstruction, segmentation, interpretation, and understanding.

Acknowledgments

We would like to thank Naomi Roberts (Dimension Studio and Unsigned Group) for permission to use Eva Herzigová MetaHuman BTS screenshots.

The paper is an output of the project NFP313010BWN6 "The implementation framework and business model of the Internet of Things, Industry 4.0 and smart transport."

Literature

Al-Emran, M., & Deveci, M. (2024). Unlocking the potential of cybersecurity behavior in the metaverse: Overview, opportunities, challenges, and future research agendas. *Technology in Society*, *77*, 102498. doi: 10.1016/j.techsoc.2024.102498

Alldieck, T., Zanfir, M., & Sminchisescu, C. (2022). Photorealistic monocular 3D reconstruction of humans wearing clothing. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, 1496-1505. doi: 10.1109/CVPR52688.2022.00156

Bertiche, H., Madadi, M., & Escalera, S. (2022). Neural cloth simulation. *ACM Transactions on Graphics (TOG)*, *41*(6), 220. doi: 10.1145/3550454.3555491

Bugaj, M., Durana, P., & Lăzăroiu, G. (2023). Virtual navigation and digital twin simulation tools, Internet of Things-based decision support and big data computing systems, and image recognition and geolocation data processing algorithms in the virtual economy of the metaverse. *Review of Contemporary Philosophy*, *21*, 137-153. doi: 10.22381/RCP2220238

Chen, L., Peng, S., & Zhou, X. (2021). Towards efficient and photorealistic 3D human reconstruction: A brief survey. *Visual Informatics*, *5*(4), 11-19. doi: 10.1016/j.visinf.2021.10.003

Correia, H. A., & Brito, J. H. (2023). 3D reconstruction of human bodies from single-view and multi-view images: A systematic review. *Computer Methods and Programs in Biomedicine*, *239*, 107620. doi: 10.1016/j.cmpb.2023.107620

Cui, L., & Liu, J. (2023). Virtual human: A comprehensive survey on academic and applications. *IEEE Access*, *11*, 123830-123845. doi: 10.1109/ACCESS.2023.3329573

Ding, Y., He, H., & Mok, P. Y. (2024). Personalizing human avatars based on realistic 3D facial reconstruction. *Multimedia Tools and Applications*. doi: 10.1007/s11042-024-19583-0

Fjørtoft Ystgaard, K., Atzori, L., Palma, D., Heegaard, P. E., Bertheussen, L. E., Jensen, M. R., & De Moor, K. (2023). Review of the theory, principles, and design requirements of humancentric Internet of Things (IoT). *Journal of Ambient Intelligence and Humanized Computing*, *14*, 2827-2859. doi: 10.1007/s12652-023-04539-3

Grupac, M., & Lăzăroiu, G. (2022). Image processing computational algorithms, sensory data mining techniques, and predictive customer analytics in the metaverse economy. *Review of Contemporary Philosophy*, *21*, 205-222. doi: 10.22381/RCP21202213

Jiang, M., Tian, Z., Yu, C., Shi, Y., Liu, L., Peng, T., Hu, X., & Yu, F. (2024). Intelligent 3D garment system of the human body based on deep spiking neural network. *Virtual Reality & Intelligent Hardware*, *6*(1), 43-55. doi: 10.1016/j.vrih.2023.07.002

Kliestik, T., Novak, A., & Lăzăroiu, G. (2022). Live shopping in the metaverse: Visual and spatial analytics, cognitive artificial intelligence techniques and algorithms, and immersive digital simulations. *Linguistic and Philosophical Investigations*, *21*, 187-202. doi: 10.22381/lpi21202212

Kliestik, T., & Lăzăroiu, G. (2023). Emotional state prediction and data visualization tools, generative artificial intelligence and workplace monitoring systems, and wearable sensor-based devices in immersive workspaces. *Analysis and Metaphysics*, *22*, 43-61. doi: 10.22381/am2220233

Lattas, A., Moschoglou, S., Ploumpis, S., Gecer, B., Deng, J., & Zafeiriou, S. (2023). FitMe: Deep photorealistic 3D morphable model avatars. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, 8629-8640. doi: 10.1109/CVPR52729.2023.00834

Lee, J., & Joo, H. (2023). Locomotion-action-manipulation: Synthesizing human-scene interactions in complex 3D environments. *IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, 9629-9640. doi: 10.1109/ICCV51070.2023.00886

Lee, J., Nguyen, D., Kim, J., Kang, J., & Lee, S. (2024). Double reverse diffusion for realistic garment reconstruction from images. *Engineering Applications of Artificial Intelligence*, *127*(B), 107404. doi: 10.1016/j.engappai.2023.107404

Liao, T., Zhang, X., Xiu, Y., Yi, H., Liu, X., Qi, G.-J., Zhang, Y., Wang, X., Zhu, X., & Lei, Z. (2023). High-fidelity clothed avatar reconstruction from a single image. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, 8662-8672. doi: 10.1109/CVPR52729.2023.00837

Liao, Z., & Waslander, S. L. (2024). Multi-view 3D object reconstruction and uncertainty modelling with neural shape prior. *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 3086-3095. doi: 10.1109/WACV57701.2024.00307

Sarakatsanos, O., PapazoglouChalikias, A., Boikou, M., Chatzilari, E., Jauk, M., Hafliger, U., Nikolopoulos, S., & Kompatsiaris, I. (2024). VR designer: Enhancing fashion showcases through immersive virtual garment fitting. *Virtual Reality*, *28*, 70. doi: 10.1007/s10055-024-00945-w

Uhlenberg, L., Derungs, A., & Amft, O. (2023). Co-simulation of human digital twins and wearable inertial sensors to analyse gait event estimation. *Frontiers in*

Bioengineering and Biotechnology, *11*, 1104000. doi: 10.3389/fbioe.2023.1104000

Valaskova, K., Vochozka, M., & Lăzăroiu, G. (2022). Immersive 3D technologies, spatial computing and visual perception algorithms, and event modeling and forecasting tools on blockchain-based metaverse platforms. *Analysis and Metaphysics*, *21*, 74-90. doi: 10.22381/am2120225

Wang, A., Xu, Y., Sarafianos, N., Maier, R., Boyer, E., Yuille, A., & Tung, T. (2024). HISR: Hybrid implicit surface representation for photorealistic 3D human reconstruction. *Proceedings of the AAAI Conference on Artificial Intelligence*, *38*(6), 5298-5308. doi: 10.1609/aaai.v38i6.28337

Xiang, D., Bagautdinov, T., Stuyck, T., Prada, F., Romero, J., Xu, W., Saito, S., Guo, J., Smith, B., Shiratori, T., Sheikh, Y., Hodgins, J., & Wu, C. (2022). Dressing avatars: Deep photorealistic appearance for physically simulated clothing. *ACM Transactions on Graphics (TOG)*, *41*(6), 222. doi: 10.1145/3550454.3555456

Yang, Z., Su, B., & Wen, J.-R. (2023). Synthesizing long-term human motions with diffusion models via coherent sampling. *Proceedings of the 31st ACM International Conference on Multimedia*, 3954-3964. doi: 10.1145/3581783.3611887

Yu, J., Alhilal, A., Zhou, T., Hui, P., & Tsang, D. H. K. (2024). Attention-based QoE-aware digital twin empowered edge computing for immersive virtual reality. *IEEE Transactions on Wireless Communications*. doi: 10.1109/TWC.2024.3380820

Zhang, J., Jiang, Z., Yang, D., Xu, H., Shi, Y., Song, G., Xu, Z., Wang, X., & Feng, J. (2023a). AvatarGen: A 3D generative model for animatable human avatars. In Karlinsky, L., Michaeli, T., & Nishino, K. (Eds.), *Computer Vision – ECCV 2022 Workshops* (pp. 668–685). Springer, Cham. doi: 10.1007/978-3-031-25066-8_39

Zhang, Z., Sun, L., Yang, Z., Chen, L., & Yang, Y. (2023b). Global-correlated 3D-decoupling transformer for clothed avatar reconstruction. *Proceedings of the 37th International Conference on Neural Information Processing Systems*, *343*, 7818-7830.

Zvarikova, K., Trnka, M., & Lăzăroiu, G. (2023). Immersive extended reality and remote sensing technologies, simulation modeling and spatial data acquisition tools, and cooperative decision and control algorithms in a real-time interoperable decentralized metaverse. *Linguistic and Philosophical Investigations*, *22*, 43-59. doi: 10.22381/lpi2220233