IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 27 May 2023, accepted 17 July 2023, date of publication 21 July 2023, date of current version 27 July 2023.
Digital Object Identifier 10.1109/ACCESS.2023.3297873

SURVEY

Intelligent Metaverse Scene Content Construction

JUNXIANG WANG[1], SIRU CHEN[1], YUXUAN LIU[1], AND RICHEN LAU[1,2]
[1]School of Computer and Electronic Information/School of Artificial Intelligence, Nanjing Normal University, Nanjing 210023, China
[2]State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China

Corresponding authors: Siru Chen (222202026@njnu.edu.cn) and Richen Lau (329789995@qq.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61702271; and in part by the Open Project Program of the State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, under Grant VRLAB2023B05.

ABSTRACT The integration of artificial intelligence (AI) and virtual reality (VR) has revolutionized research across various scientific fields, with AI-driven VR simulations finding applications in education, healthcare, and entertainment. However, existing literature lacks a comprehensive investigation that systematically summarizes the fundamental characteristics and development trajectory of AI-generated visual content in the metaverse. This survey focuses on intelligent metaverse scene content construction, aiming to address this gap by exploring the application of AI in content generation. It investigates scene content generation, simulation biology, personalized content, and intelligent agents. Analyzing the current state and identifying common features, this survey provides a detailed description of methods for constructing intelligent metaverse scenes. The primary contribution is a comprehensive analysis of the current landscape of intelligent visual content production in the metaverse, highlighting emerging trends. The discussion on methods for constructing intelligent scene content in the metaverse suggests that in the era of intelligence, it has the potential to become the dominant approach for content creation in metaverse scenes.

INDEX TERMS Content generation, metaverse, immersive visualization, deep learning.

I. INTRODUCTION

The rapid evolution of metaverse technologies, coupled with exponential growth in computing power, has brought about a paradigm shift in the production of visual content. The traditional manual creation of static scene content has been revolutionized by the emergence of intelligent metaverse technology, which enables the generation of scene content with unprecedented capabilities. This intelligent approach leverages existing data conditions as specialized tools for programmatically constructing and enhancing metaverse scenarios. The application of intelligent metaverse scene content generation extends to various scientific domains including education, biology, medicine, and art. Specifically, it involves the dynamic construction of multi-variable scenes through programming.

In contrast to conventional artificial intelligence-generated content (AIGC), which predominantly focuses on generating text, images, audio, video, and models, our research focuses on the utilization of AI methods to generate content within metaverse scenes. Distinct from AIGC's video content generation [1], our emphasis is on the creation of specific scenes that empower users to engage in free interaction within the generated metaverse environment. This essential feature sets our content generation approach apart from simple videos, as it fosters a highly immersive and interactive user experience integrated with the interaction concept of Metaverse [2]. Based on the aforementioned considerations, we have compiled and analyzed recent research on the construction of intelligent metaverse scene content. We have classified, summarized, and discussed these works in order to inspire future endeavors in this field.

The rapid advancement of AI technologies has led to the development of various methodological models that integrate AI intelligence into automated metaverse scene generation. Three primary strategies have emerged: the development of intelligent framework systems for coordinating metaverse content, simulation of intelligent agents that replicate

The associate editor coordinating the review of this manuscript and approving it for publication was Xiaogang Jin.

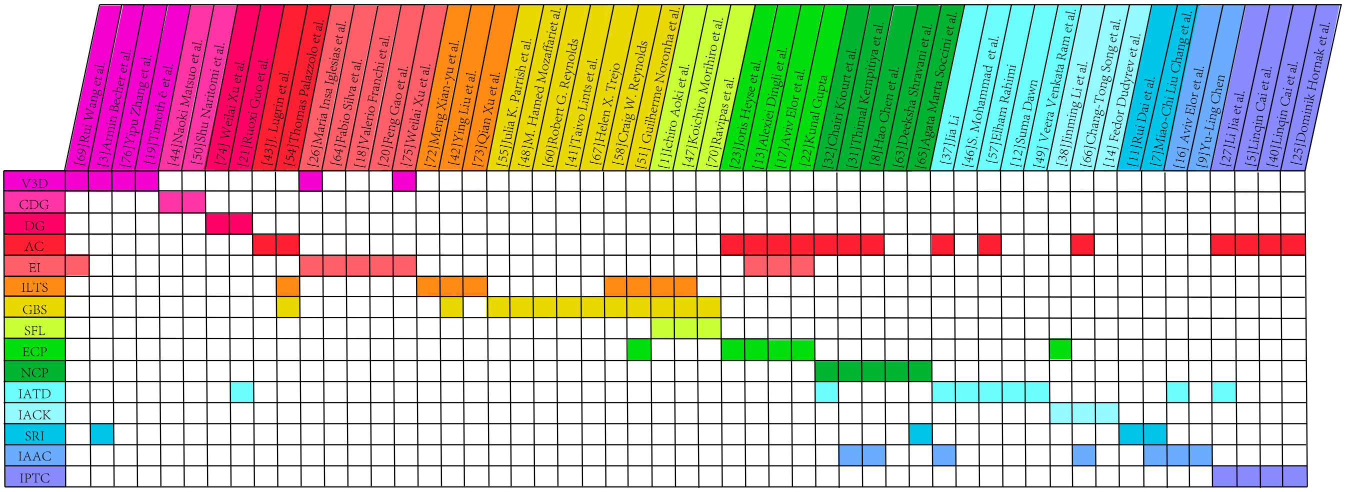


FIGURE 1. We choose a few exemplary papers from each category. Different publications are shown on the horizontal axis, while various kinds of intelligent visual content production strategies are represented on the vertical axis, including scene content generation (V3D: virtual 3D construction, CDG: classification-based directed generation, DG: data-driven generation, content construction and improvement (AC: adaptive change, EI: effect improvement)), simulated biology (ILTS: individual life trait simulation, GBS: group behavior simulation, SFL: stress feedback learning), personalized contents (ECP: emotional content personalization, NCP: non-emotional content personalization), intelligent agents (IATD: intelligent agents trained based on data, IACK: intelligent agents constructed from external knowledge systems, intelligent agents trained for interaction data adaptation (SRI: simulating realistic images, IAAC: interaction actions to adapt agent changes, IPTC: interaction processes to adapt task collaboration)). Many of the papers share the basic characteristics of multiple intelligent visual content generation techniques at the same time. All the papers are categorized (in different colors) and disaggregated according to the main techniques they use.

biological logic, and utilization of machine learning or deep learning methods to generate or enhance scene content. Understanding the comprehensive integration of intelligent metaverse scene content generation requires a detailed study of these strategies and an analysis of the employed methods to reveal valuable insights into the future trends of metaverse scene content generation in the era of AI.

A. RELATED SURVEYS

Previous surveys in this field have predominantly focused on AI-based computer vision imaging, with particular emphasis on evolutionary computation and biological vision [3], or on generating interactive and immersive content [4]. Some surveys have delved into the impact of content generation in AI on the art industry [5] and explored deep learning-based image style transfer [6]. Additionally, the Unreal 3D authoring engine has shown interest in leveraging agents and decision-tree-based procedural content generation techniques for immersive game development [7].

In comparison to studies that delve into the application of AI technology in machine learning for character programming and calculation in games [8], or compare search-based methods and machine learning methods to traditional approaches in surveying game program content [9], our survey takes a more comprehensive approach by focusing on the application of AI technology and the generation of intelligent metaverse scene content. While there have been surveys that explore the inclusion of human emotions in virtual agents and the advancements of AI in education [10], these surveys either have limited coverage of content generation or primarily focus on detailed descriptions of a few representative studies [11], [12], [13]. In contrast, our survey offers a comprehensive analysis of recent advancements from a cutting-edge perspective, incorporating relevant papers from the IEEE Xplore and ACM Digital Library based on their relevance and research direction.

B. TAXONOMY OF THE SURVEY

The exploration of AI for generating content in metaverse scenes has led to the emergence of several common areas of interest, including scene content generation, simulated biology, personalized content, and intelligent agents, either individually or in combination (Figure 1). It is important to highlight that many methods employ a combination of these intelligent metaverse scene content generation techniques rather than relying solely on a single approach. While this survey does not encompass all papers related to metaverse scene content generation, its primary focus is on the utilization of deep learning or other machine learning methods and intelligent frameworks to achieve the generation of metaverse scene content.

II. SCENE CONTENT GENERATION

Integrating AI into the design framework of VR presents a logical solution for overcoming the challenges associated with automated metaverse scene content generation, thereby enhancing the realism of virtual environments. Deep-learning modules have emerged as effective tools for achieving this objective. Notably, Unreal Engine explored AI-based approaches by utilizing agents and decision trees to generate content in immersive game programs [7].

This section introduces several common intelligent methods used in the generation and improvement of metaverse scene content. These methods can be broadly divided into four main categories: virtual 3D reconstruction, directed generation based on classification, data-driven generation, and content construction and improvement. It is worth mentioning that certain approaches may exhibit characteristics that align with multiple categories simultaneously. For this study, we classified them based on their primary features.

A. VIRTUAL 3D CONSTRUCTION

Two-dimensional photos are crucial for reconstructing three-dimensional scenes because of their accessibility and the availability of comprehensive data. Wang et al. proposed a method that utilizes the ant colony algorithm to rapidly segment video scenes and combines the segmentation process with the similarity of the same modal data and the correlation of different modal data to construct multiple panoramic-image arrangements, thus creating a VR scene space [14]. Similarly, VIRTOOAIR significantly improves the system's posture recovery capability by incorporating VR tracking data and low-semantic input from RGB photos into an end-to-end reconstruction process using deep learning algorithms, as shown in Figure 2(a) [15]. In contrast to purely two-dimensional image data, Zhang et al. applied effect processing to two-dimensional images, as shown in Figure 2(c). The 2D input images were transformed using a visual art language with a traditional Chinese ink and wash style, followed by a 3D virtual technique to generate a 3D virtual scene in ink and wash style [16]. Freville et al., on the other hand, indirectly employed YOLO for object detection to identify realistic objects and generate corresponding 3D world assets at their respective locations, as depicted in Figure 2(b) [17].

These approaches demonstrate the potential of integrating deep learning techniques and innovative algorithms for constructing metaverse scenes. However, challenges remain in terms of handling complex scenes, ensuring the accuracy of the reconstruction process, and achieving a balance between realism and efficiency. Further research is required to address these challenges and advance the field of virtual 3D construction in metaverse environments.

B. CLASSIFICATION-BASED DIRECTED GENERATION

Generating different categories of objects often requires specialized treatment to achieve the desired results. Naoki Matsuo et al. employed deep learning networks to classify objects into "spatial boundaries" and "common obstacles." By overlaying virtual objects onto real objects based on their classifications, they created a more recognizable VR space that remained enjoyable even in cluttered environments filled with obstacles [18], as depicted in Figure 3(a).

Similarly, the generation of food textures can be expedited through the pre-categorization of food items. The FoodChangeLens approach [19] uses a cyclic generative adversarial network (GAN) trained on a large-scale collection of food images from Twitter streams. This enabled the

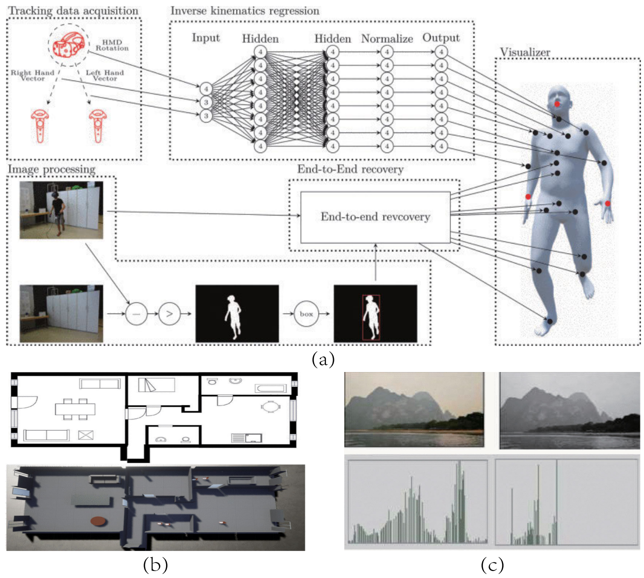


FIGURE 2. Virtual 3D construction. (a) The processing pipeline for VIRTOOAIR. The data-collecting module for VR tracking is the initial system component. To learn the upper body joint configuration, a deep neural network is given the VR tracking data (i.e., HMD and hand motions). Inverse kinematic regression of the deep neural network is performed using the second module. Processing of RGB images is handled by the third module, which also has a pre-processing procedure (i.e. extraction of bounding boxes). To rebuild the lower body joint rotation, the bounding box and camera pictures are both continually input into the end-to-end recovery framework (i.e., the fourth module). The final module's visualization tool is in charge of properly portraying the virtual body in VR [15]. (b) On a practice floor plan, tests were conducted. The original floor layout can be seen in the top image, while the 3D environment can be seen in the bottom image [17]. (c) Engaging AI [16].

transformation of food categories while preserving the shape of the original food item, as illustrated in Figure 3(b). This method leverages conditional CycleGAN (cCy-cleGAN) [20] to perform image transformations, along with food segmentation results obtained from U-Net [21], a convolutional network designed for biomedical image segmentation. The transformed images were rendered as textures.

These classification-based directed generation methods offer a more targeted approach to scene generation and demonstrated the potential of deep learning techniques to improve the realism and efficiency of metaverse content creation. However, challenges remain, such as the need for comprehensive and diverse training datasets, and the potential for bias in classification systems. Further research is needed to address these issues and to develop more sophisticated and reliable classification-based generation methods.

C. DATA-DRIVEN GENERATION

Data-driven generation relies on diverse collections of realistic data to achieve accurate and immersive simulations of various scenarios. For example, in the field of football, a trained PFNN model [22] can be used to make predictions and create a fully immersive 3D environment, allowing analysts, coaches, and players to view an entire game. By capturing and learning human flow patterns through agent trajectories powered by a neural-network-based animation system, player motions in the field and their dynamics from one event to the next can be accurately represented [23], as depicted in Figure 3(c). In the domain of music instruction, hand motions can be automatically generated for 3D animations using MIDI data as the input. This is accomplished by leveraging a pre-trained Hidden Markov Model (HMM) for fingerings detection [24], as shown in Figure 3(d).

Data-driven generation techniques present exciting opportunities for creating realistic simulations and instructional metaverse animations. The integration of diverse data sources, such as football trajectories and MIDI inputs, enables immersive experience and skill development. However, addressing the challenges of data collection, dataset quality, and fostering creativity in content generation are important areas for future research. By addressing these limitations, researchers can unlock the full potential of data-driven generation in metaverse environments and further enhance immersive experiences and interactive capabilities.

D. CONTENT CONSTRUCTION AND IMPROVEMENT

The construction and adaptation of content are essential for creating engaging scenarios that provide users with unique experiences. To achieve this, content must be generated in a manner that fits realistic scenarios and provides natural variability in the development of different scenarios. Researchers have explored two approaches to achieving this: adaptive changes to content and effect improvements.

1) ADAPTIVE CHANGE

For an adaptive change, Lugrin et al. used an AI module to generate event chains from a triggering perspective [25], as shown in Figure 4(a). They simulated the spontaneous movements of objects, causing collisions between them, and selected modifications closest to a predefined cost threshold based on heuristic values. This approach triggers conceptual events related to the user experience and results in content that adapts to the user's actions and choices.

Another example of adaptive change is DeepDive [26], which incorporates AI tools to simulate spontaneous changes in the environment to recognize the patterns of systems in the resource environment of ancient civilizations, as shown in Figure 4(b). This regional growth approach uses pattern recognition in hydrology, AI collection modules on vegetation to predict potential vegetation of cells, A*, and cultural algorithms to generate reindeer herds, providing opportunities to generate and test anthropological and archaeological theories.

These adaptive change approaches offer users a unique and dynamic experience that adapts to their choices and actions, thereby creating more engaging scenarios.

2) EFFECT IMPROVEMENT

To improve the effectiveness of generated content, researchers have focused on enhancing the visual and delivery aspects of content. Iglesias et al. proposed feature extraction using AI algorithms [27], as shown in Figure 4(c). They employed visualization methods and new UE4 classes to register and superimpose the collected characteristics on a 3D model, enabling the identification of structural flaws using an AI-based approach. This technique improves the photorealism and immersive qualities of content.

QM4VR [28], on the other hand, focuses on delivery improvements rather than 3D model upgrades. By monitoring the operation and performance of the sub-streams and modifying the load balancing/scheduling method to enhance the quality of service (QoS) parameters, the VR Content Quality Multipath Delivery (QM4VR) approach manages MPTCP sub-streams to distribute priority packets based on content awareness, as shown in Figure 4(d). This enhances the delivery of VR content and ensures a smoother and more reliable user experience.

Furthermore, the rover positioning system [29] offers a novel approach for improving the overall effect of the VR experience, as depicted in Figure 5(a). By utilizing a bird's eye view of the VR world and employing the MCL algorithm, monocular visual-inertial odometry, and Siamese neural networks [30], the system predicts the rover's position by adding GAN-enhanced images to the VR image dataset, and the rover's performance in the VR environment is enhanced.

To enhance the details of the scene contents, Feng Gao et al. applied convolutional neural networks in AI technologies to extract and reconstruct a virtual field point cloud [31], as shown in Figure 5(b). They established a coordinate matching relationship between a computer-generated virtual scene and actual reality and improved the scene's details using techniques such as the Gouraud shading algorithm and texture mapping. This approach reduces distortion and enhances user satisfaction by providing a more visually appealing and realistic virtual environment.

To address the challenge of sustainable scene style transformation in real-time, Lei Yang et al. proposed methods to increase the frame rate of real-time rendering applications and applied them to the Barracuda style transformation in VR [32], as shown in Figure 5(c). By optimizing the rendering process, they enabled smooth and consistent scene style transformations, improving the overall visual quality and coherence of the VR experience.

These approaches focus on enhancing the visual quality, delivery efficiency, and overall immersion in metaverse content, leading to more engaging and satisfying user experiences.

III. SIMULATED BIOLOGY

Simulating real-life situations through the generation of virtual content has emerged as a valuable approach in the biological sciences. It offers researchers the opportunity to study and analyze complex biological systems, providing insights that may be challenging to obtain through direct observation alone. In recent years, there has been a growing focus on utilizing AI-generated content to simulate the social attributes

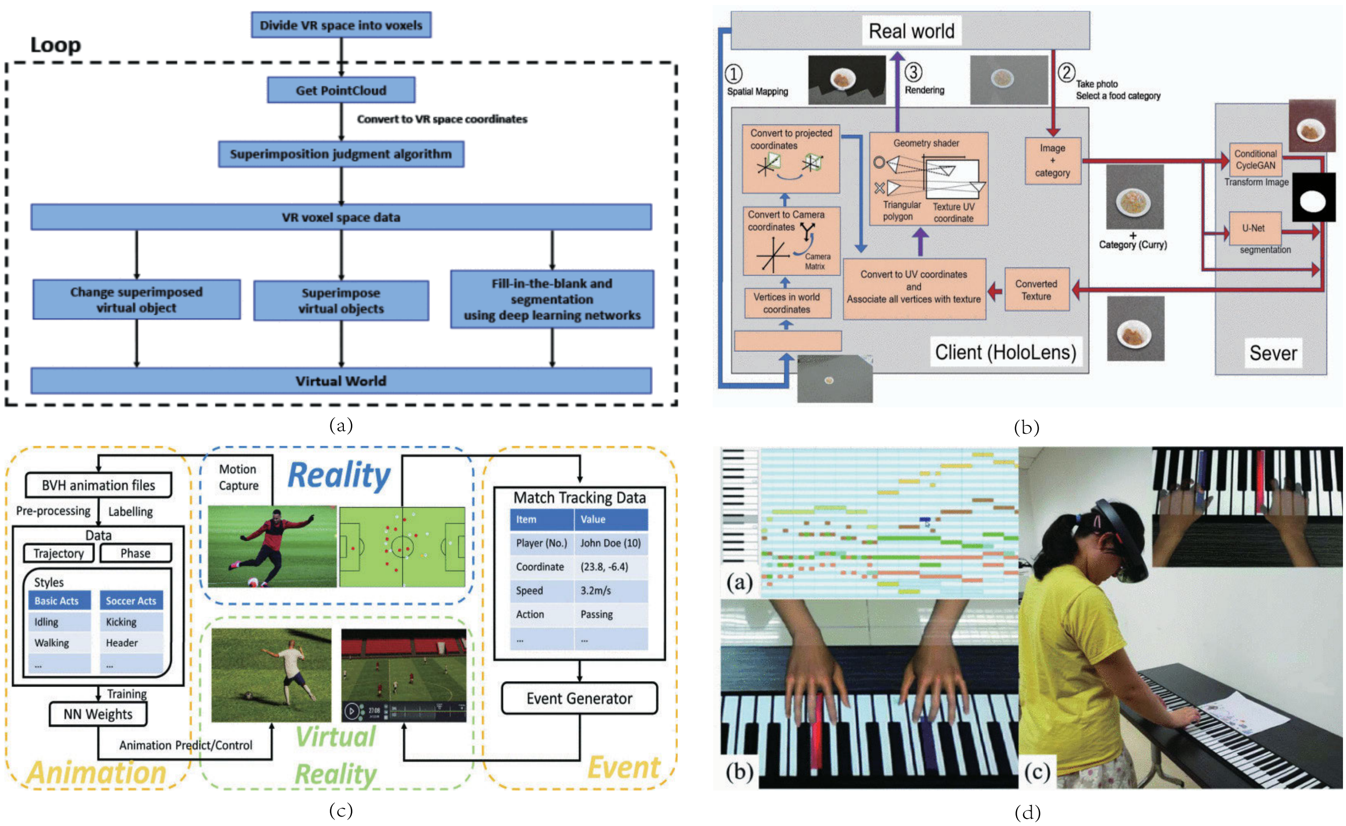


FIGURE 3. Classification-based directed generation. (a) Real-time method processing pipeline for creating VR environments based on reality [18]. (b) Overview of the FoodChangeLens holographic system [19]. (c) Overview of the system: Motion-captured animations of individual football players are utilized to create virtual agents (football players), and data taken from recorded football matches is used to create replay events in the VR simulation [23]. (d) The input MIDI file comprises time sequences for the various keys, as seen in the top left image. Bottom left image: Our technology creates a 3D performance animation for both hands based on the MIDI file. The figure on the right shows how the performance animation is displayed in the HMD and is in tune with a real piano, giving the user real-time visual cues and feedback [24].

and habits of various organisms, including fish, flocks, and birds, within biological ecological environments [33], [34].

This section discusses the application of AI content generation in simulated biology, covering three crucial aspects: individual life trait simulation, group behavior simulation, and stress feedback learning. By exploring these areas, researchers aim to gain a deeper understanding of the behavior and dynamics of biological systems in virtual environments.

A. INDIVIDUAL LIFE TRAIT SIMULATION

The simulation of individual organisms is crucial for simulating living organisms and for enabling the exploration of ecological dynamics. Simulating the characteristics of different organisms requires the application of fundamental principles. For instance, the simulation of an artificial fish captures the essence of simulating individual organisms. To achieve a more realistic simulation of individual fish lives, researchers have proposed various methods.

Xian-Yu et al. developed a multi-sensory system model that incorporates vision, touch, taste, and smell to mimic the routines and behavior of fish in the wild, thereby enhancing the simulation of the physiological characteristics and individual responses of artificial fish [35], as shown in Figure 6(b). To simulate diverse variations in the process of fish egg division, Xu and Zhou suggested a fish egg model based on the artificial life approach, which incorporates life dynamics, artificial life methodology, AI, and genetic algorithms [36], as shown in Figure 6(a).

However, purely artificial life approaches have inherent time limitations. To address this, Liu proposed the integration of an evolutionary model into the artificial fish reproduction process, enabling the gradual enrichment of the vital signs of fish over time through an overall evolutionary performance [37], as shown in Figure 6(c). This approach has the potential to generate more realistic simulations of individual organisms by modeling gradual changes in their life traits over time.

The limitations of these approaches are not explicitly mentioned, but the integration of an evolutionary model to overcome time limitations suggests that the time-consuming nature of simulating individual organisms is a challenge. Additionally, although these approaches may improve the simulation of individual organisms, there may still be limitations in accurately simulating complex ecological dynamics. Further research should focus on addressing these

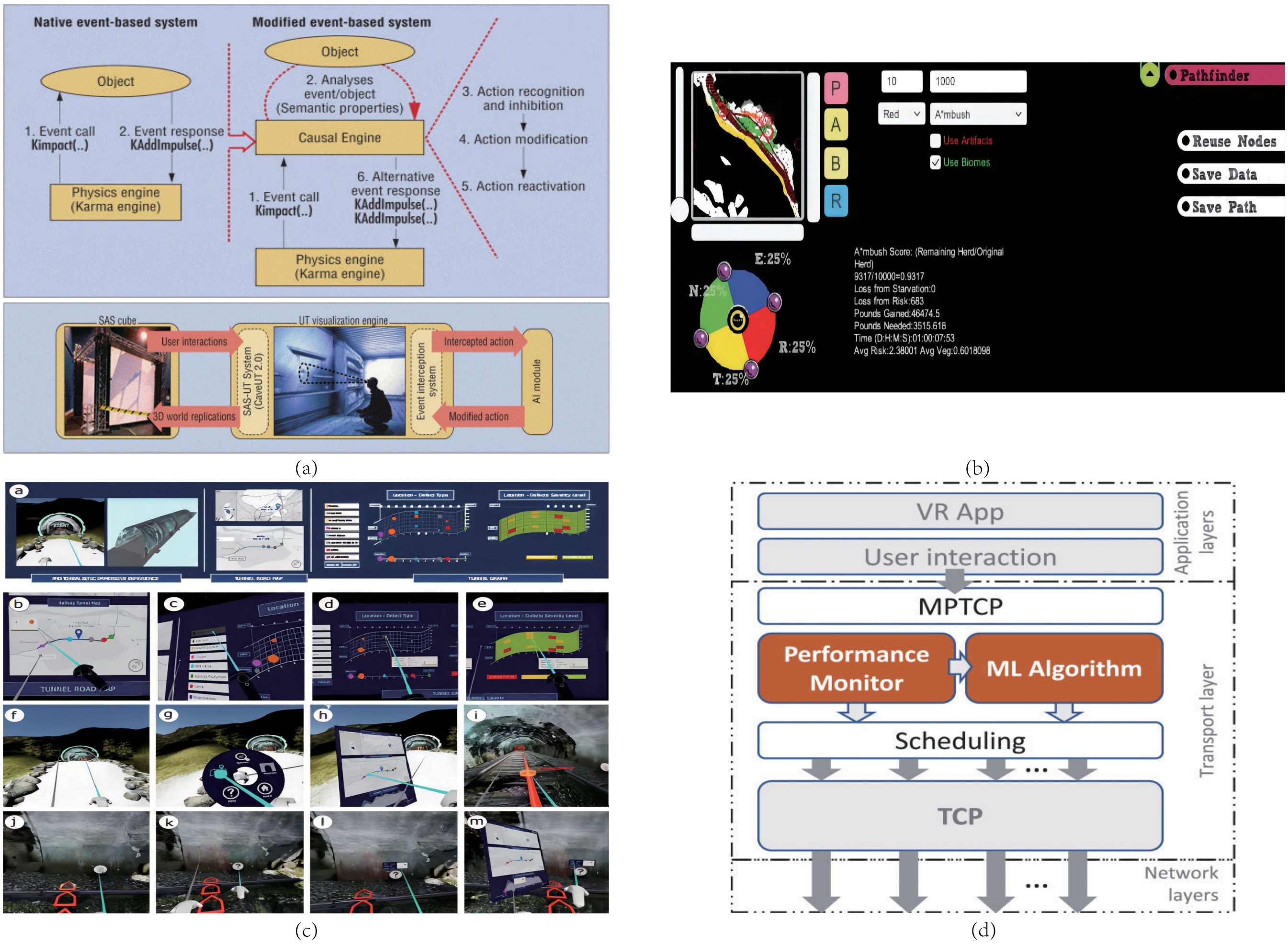


FIGURE 4. Content construction and improvement. (a) AI-based behavior has been included in the game engine's event system, as seen in the top side figure. View of the system architecture in the bottom figure. The system is built on a gaming engine that has been converted to an immersive display similar to CAV [25]. (b) Screenshot of the in-depth exploring Interface for DeepDive. An A* ambush herd's course for a specific herd size is shown on a map in the upper left corner. Based on the existing rule set, the red circles on the map represent potential site locations. The wheels below show the algorithm's herd priority as determined by the user (N = nutrition; R = risk; E = movement force; and T = time to target) [26]. (c) The dashboard (a-e) of the EPI system is immersive, while the sceneries (f-m) depict real-world surroundings [27]. (d) The fundamental design of QM4VR [28].

limitations to enhance the accuracy of ecological simulations and promote a better understanding of ecological systems.

B. GROUP BEHAVIOR SIMULATION

Simulating group behavior is a critical aspect in creating a comprehensive representation of ecological dynamics, and the integration of AI technologies has revolutionized the vividness of these simulations in VR environments. Researchers have proposed various innovative approaches to tackle the challenges associated with simulating group behavior.

One notable approach is the use of individual-based cellular automata (CA) models [38] to refine the understanding of schooling behavior and guide the construction of natural collective behavior by specifying individual-level rules. This method not only provides a realistic representation of group dynamics but also allows for the exploration of emergent properties that arise from interactions between individuals. For instance, Mozaffari et al. introduced a machine learning tracking technique that employs a Bèzier curve method to smooth the movement of fruit flies, resulting in more natural and lifelike behavior. Additionally, logistic regression classification was utilized to classify the mating state of the fish based on video frames captured before training [39], as shown in Figure 7(a). This innovative approach enhances the realism of group behavior and enables a more accurate representation of ecological systems.

Although holistic simulations are valuable, they often lack ecological fluidity. To address this limitation, researchers have proposed integrating external stimulus demands and adaptive mechanisms into simulations of artificial life. Lints developed a cluster scheduling algorithm that considers the characteristics of both individuals and groups, allowing for integrated judgment of artificial life and better adaptation to environmental changes [40], as shown in Figure 7(b). Moreover, the incorporation of artificial evolution in group

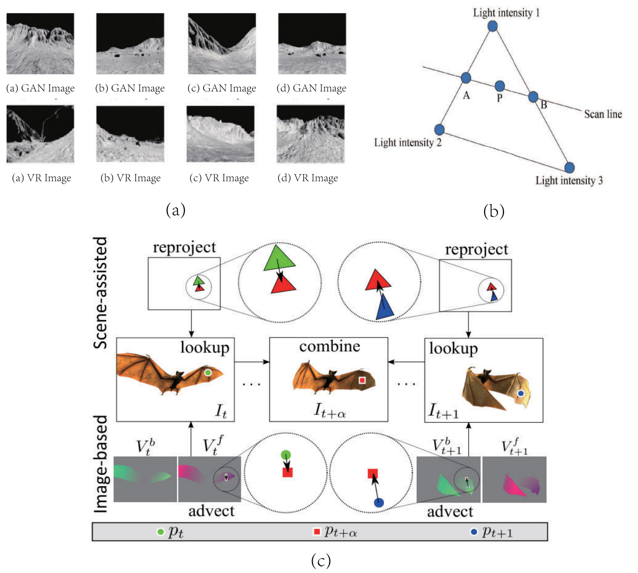


FIGURE 5. **Effect improvement. (a) The picture produced by a GAN and a VR image [29]. (b) Diagram of the Gulo coloring algorithm's processing of light intensity [31]. (c) Overview of the algorithm: pixels $I_{t+\alpha}[p_{t+\alpha}]$ in frame B are recreated from consecutive frames I using scene-aided and image-based reprojection. Whereas image-based reprojection estimates streams by repeatedly exploring nearby frames for depth and 3D flow fields, scene-aided projection generates scene streams during conventional triangle rasterization [32].**

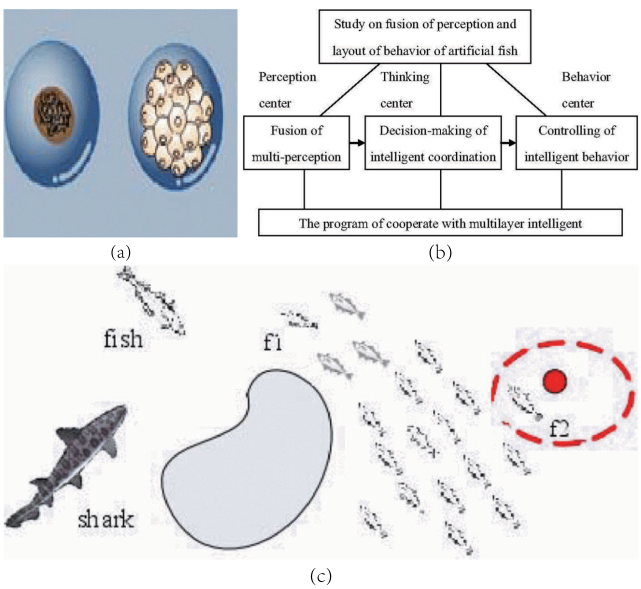


FIGURE 6. **Individual life trait simulation. (a) The process by which fish eggs divide their cells [36]. (b) The connection between perception, choice, and behavior in synthetic fish [35]. (c) An artificial fish is depicted in a schematic [37].**

simulation, as exemplified by the adjustment of boid parameters during evolution [41], facilitates innovation in the field of group intelligence. This innovative approach enables the exploration of novel group behaviors and the emergence of collective intelligence in simulated ecological systems.

To address the movement challenges and enhance the realism of VR environments, researchers have proposed simulating crowd behavior while considering obstacle avoidance. The walking areas of crowds can be determined using an A* search algorithm, resulting in more natural movement patterns [42], as shown in Figure 7(c). Moreover, the coordination of group movements was achieved through the application of a multi-intelligent distributed system model, as proposed by Reynolds et al. [41]. This innovative approach ensures that group behavior remains cohesive and avoids chaotic movements, thereby enhancing the overall realism of the simulation.

Furthermore, the influence of emotions on group behavior was explored to add a layer of realism. Noronha et al. introduced the impact of panic emotions, which modifies the behavioral parameters of agents and influences their decision-making processes, particularly regarding the search for an ideal path within a crowd [43], as shown in Figure 7(d). This innovative integration of emotional dynamics enhances the believability of group behavior and contributes to a more immersive simulation experience.

The application of AI technologies in group behavior simulation within ecological contexts has resulted in notable advancements, including enhanced realism, improved adaptability, and incorporation of complex dynamics. These innovations have expanded our understanding of collective behavior in ecological systems and contributed to the emergence of new insights and the development of intelligent simulations. However, it is important to acknowledge that there may still be limitations and challenges in accurately representing the intricacies of real-world ecosystems. Scalability, computational demands, and the ability to capture the full complexity of ecological dynamics are potential areas for further research and improvement. Addressing these limitations can help refine the application of AI in group behavior simulations, ultimately advancing our understanding of ecological phenomena and enhancing the effectiveness of intelligent simulations.

C. STRESS FEEDBACK LEARNING

Behavior in natural ecosystems is not static or confined but responds dynamically to various stressors. To accurately simulate ecological systems, it is crucial to incorporate stress feedback learning, which enables organisms to adapt and evolve in response to unknown situations, cooperative interactions, competition, and other biological relationships. This section explores innovative approaches to stress feedback learning in the context of simulated biology.

In the realm of fish populations, the incorporation of a simplified and evolving model of fishing gear can stimulate sportive evolution through predation dynamics [44], as depicted in Figure 8(a). This approach introduces selective pressures that drive the adaptation and survival of fish populations, mimicking the real-world dynamics of predator-prey relationships and environmental changes. By incorporating such stress factors, the simulation becomes more realistic and allows the observation of evolutionary responses within a virtual ecosystem.

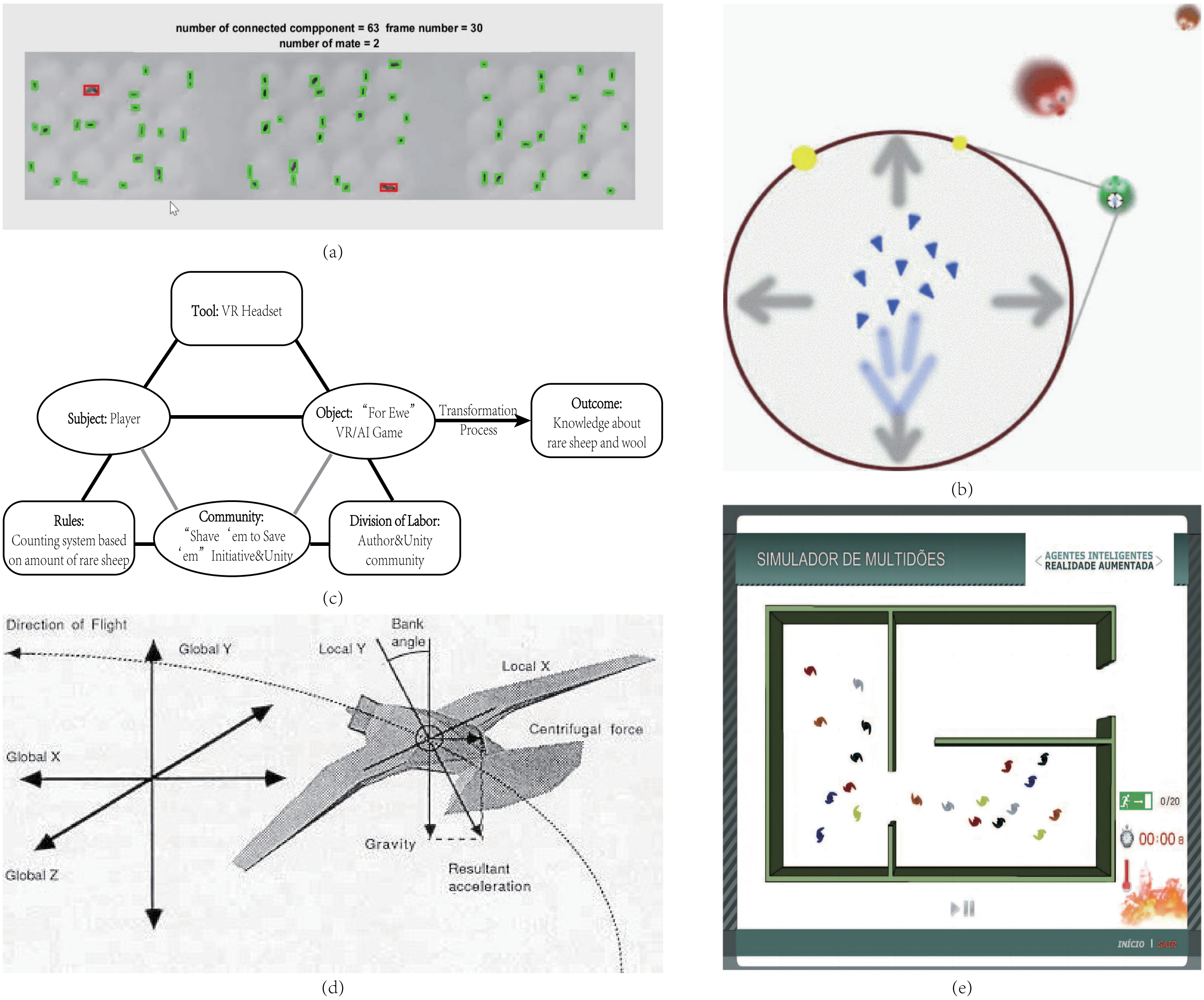


FIGURE 7. Group behavior simulation. (a) An image processing-based one-frame system. Fruit fly identification is shown by the green frame, and mating is indicated by the red frame [39]. (b) A straightforward flock-based defense mechanism to keep 'enemy' at bay. Our controlled creature (facing up) conflicts with two evil entities in the image's upper right corner (facing downwards). The enormous circle on the left is a virtual space that serves as our creature's "brain"; in other words, it may be seen as being housed inside the creature's head rather than next to it as is shown for visualization. The room is filled with a collection of little triangles. There are two 'lights' (circles) that stand in for the adversary on the room's walls. There are no little bodies around the lights (large sketch arrows pointing down). The controlled creature will begin to travel backward when the flock reaches the base of the circle (the four smaller grey arrows indicate the mapping between the position of the flock and the control signal sent to the creature) [40]. (c) Activity theory is being used to direct the VR/AI game 'For Ewe' [42]. (d) An illustration of a mock bird [41]. (e) Zuangentes test screenshot [43].

Another approach to stress feedback learning involves the incorporation of predator-prey relationships into self-organizing adaptive schemes. Using reinforcement learning techniques [45], a population of agents can learn and adapt through trial-and-error exploration by adopting different behavioral traits when encountering similar agents and predators to maximize rewards from the environment [46], as shown in Figure 8(b). This innovative approach allows for the emergence of complex behaviors and strategies within the simulated population, providing insights into the dynamics of predator-prey interactions and the evolution of adaptive responses.

Incorporating additional factors such as predators and food sources further enhances the realism of the simulation. By applying reinforcement learning techniques to a school of fish, a simulation model can train ML agents using deep reinforcement learning algorithms [47], as depicted in Figure 8(c). This approach enables fish to learn and adapt their behaviors in response to changing environmental conditions, predators, and food availability. Through stress feedback learning, the simulated fish population exhibited sophisticated and realistic responses, closely mirroring the dynamics observed in natural ecosystems.

These innovative applications of stress feedback learning in simulated biology offer valuable insights into the adaptive capabilities of organisms and their response to changing environments. These simulations capture the dynamics of biological systems by simulating stressors and employing learning algorithms, thereby enabling the study of emergent properties and evolutionary processes. This approach provides researchers with a deeper understanding of the complex interactions and behaviors exhibited by organisms when exposed to different stressors. Nevertheless, it is crucial to acknowledge the limitations of these methods. Challenges may include accurately representing the full complexity of real biological systems, scaling simulations to larger and more intricate scenarios, and ensuring the realism and accuracy of stress feedback mechanisms. Overcoming these limitations will pave the way for further advancements in simulated biology and will contribute to a more comprehensive understanding of the natural world.

IV. PERSONALIZED CONTENTS

The quality of artificially intelligent content generation in a VR environment is determined by both the generation capabilities of the system and the input it receives. When generating intelligent content, it is crucial to ensure that the generated content meets the needs of real users, as this directly affects the VR experience and user satisfaction. Therefore, it is essential to improve the outcome of generated content based on the specific needs of individual users.

To address this challenge, several solutions for personalized content generation have been developed. These solutions can be broadly categorized into two groups: those that capture sentiment data for analytical feedback to make the content more personalized, and those that generate content adjusted by non-emotional data feedback from within VR to cater to the user's preferences.

The first category focuses on capturing sentiment data, such as user emotions and preferences, to provide analytical feedback. By analyzing this feedback, the system can gain insight into the user's emotional state and tailor the generated content accordingly. This approach ensures that the content resonates with the user's emotions, thereby creating a more personalized and engaging VR experience.

The second category involves generating content that is adjusted based on non-emotional data feedback obtained from within the VR environment. This feedback summarizes the user's tendencies and preferences, allowing the system to generate content that aligns with the user's specific needs. By adapting the content to the user's preferences, this approach enhances user satisfaction and the overall enjoyment of the VR experience.

By incorporating personalized content-generation solutions, VR environments can provide users with tailored experiences that cater to their individual needs and preferences. This not only enhances user satisfaction but also maximizes the potential of artificially intelligent content generation in delivering impactful and engaging VR experiences.

A. EMOTIONAL CONTENT PERSONALIZATION

Emotions play a crucial role in subjectively evaluating the quality of immersive content. Incorporating personalized sentiment as a target of intelligent learning can effectively fine-tune the results of online content generation, ensuring that the generated content is highly relevant to the user's emotional state. This approach enhances the overall user experience by delivering emotionally engaging and tailored content.

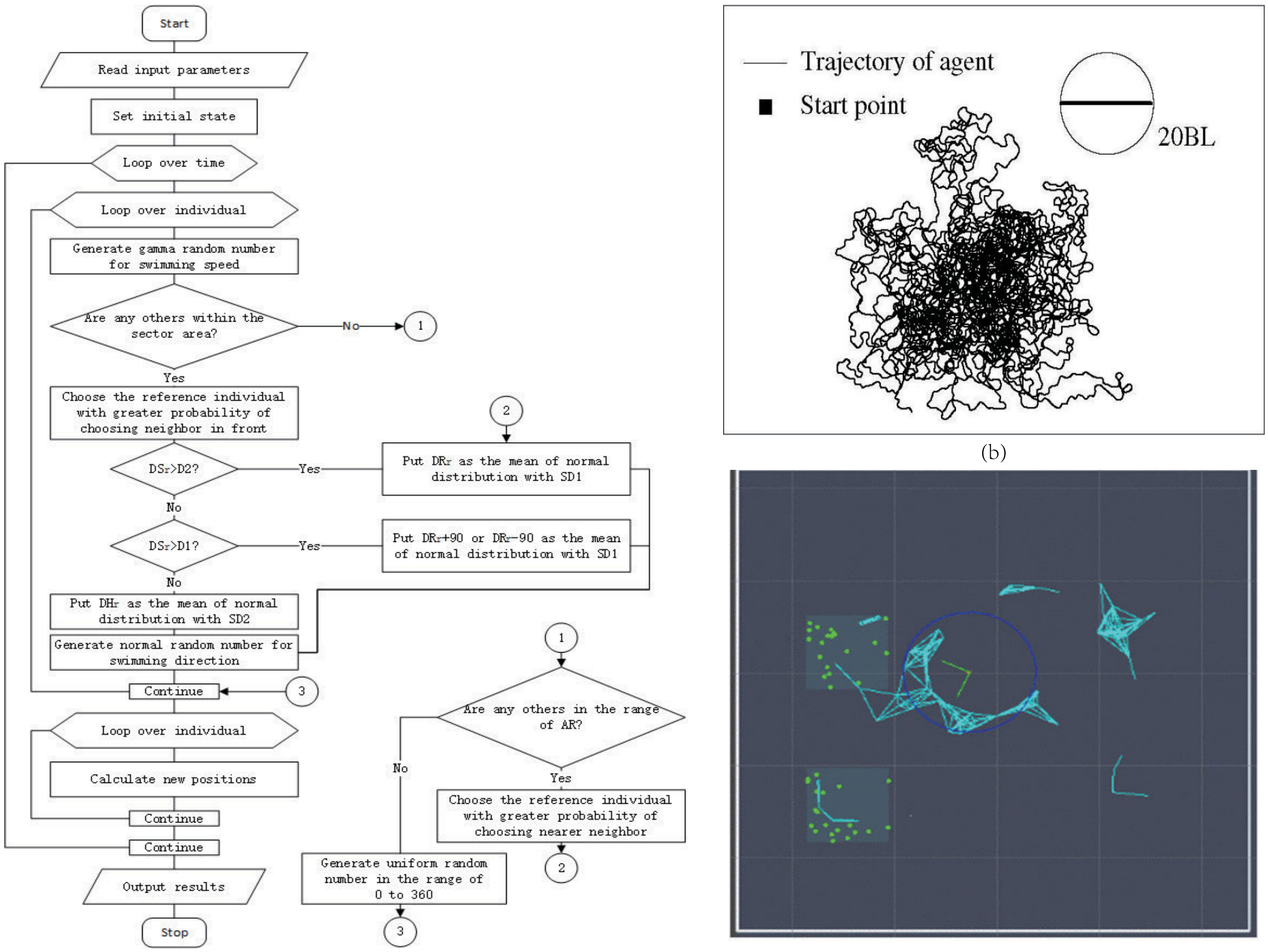
An example of personalized VR content is VRelax, a personalized VR relaxation therapy approach proposed by Heyse et al. [48], as shown in Figure 9(a). VRelax utilizes semantic reasoning and a decision-making approach based on multi-armed bandit (MAB) reinforcement learning. It leverages static and dynamic personality data as inputs to select environmental decisions adaptively. By personalizing the virtual environment and therapy experience, VRelax improves relaxation therapy outcomes in individuals with mental health issues.

The ISAM model [49] is another approach that enhances emotional content personalization. This model utilizes ten practice images from the International Affective Picture System (IAPs) and a collection of images from Google Images to improve mood prediction based on the pleasure-arousal-dominance (PAD) model [50], as shown in Figure 9(b). By intelligently selecting pictures based on the PAD reaction of the user, the ISAM model tailors the content to the emotional state of the user, enhancing the emotional impact of the VR experience.

Dingli and Bondin adopted a different approach by collecting data through wearable devices and using adversarial generative networks (GAN) to predict a user's emotional journey [51], as shown in Figure 9(c). This allows them to link emotional states to specific events in the VR experience, enabling the modeling of changes and game adaptations based on the user's emotional responses to factors such as color and sound. This personalized approach enhances emotional resonance and immersion in a VR environment.

By contrast, the PerAffect-lyVR system proposed by Gupta focuses on training with response information and leveraging situational and physiological cues from the user's interaction to detect emotional states [52], as shown in Figure 9(d). By adapting to VR environments and difficulty levels based on the user's emotional state, PerAffect-lyVR provides a highly personalized and emotionally engaging experience.

By incorporating emotional content personalization techniques, VR systems can offer tailored experiences that align with user emotions and preferences. However, it is important to address the ethical concerns associated with the use of personal emotional data in virtual environments. Despite these considerations, these approaches contribute to enhanced emotional engagement, immersion, and overall satisfaction, resulting in more impactful and meaningful metaverse experiences for users. Advancements in intelligent methods for recognizing human emotions [53] further fuel the



(a) (c)

FIGURE 8. **Stress feedback learning. (a) The simulation model's flowchart. DS_r represents the separation between the reference and the specified individuals, and DR_r represents the direction in which they are separated. DH_r, name of the reference; AR stands for an angular range of interactions. D1, the avoidance range; D2, the closeness; SD1, the standard deviation of movement in the avoidance and approach directions; SD2, motion in the parallel direction [44]. (b) Agent trajectory learning in the 0-500 step range without predators N=10, (R1, R2, R3) = (4, 20, 50) [46]. (c) The perceived radius of the predator (dark blue circle), the line linking the fish agents to one another (light blue line), and the distance to the meal are all displayed by a dense school of trained models on all behaviors (solid green circle) [47].**

potential of combining personalized content technology and emotions, unlocking new possibilities in the field.

B. NON-EMOTIONAL CONTENT PERSONALIZATION

In addition to emotions, objective data generated from VR content interactions can provide valuable insights into personalized content generation. These more deterministic data, allow for effective profiling of the user experience, enabling the generation of personalized content based on specific preferences and needs.

Kiourt et al. designed a personal virtual museum (PVM) that incorporates dynamic systems and pan-institutional modular learning objects to facilitate collaboration, knowledge modeling, and management [54], as shown in Figure 10(a). Their approach utilizes tree search, reinforcement learning, supervised learning, and a wide range of computational, audio, video, and image-based media. By leveraging structured external connections, the PVM enhances the contextual relevance of the virtual world, delivering dynamic, enjoyable, and engaging experiences tailored to the user's needs.

Personalized game experiences can be achieved by adjusting the difficulty level and utilizing personalized profiles generated from each player's Hyperseed [59]. Additionally, customized categorization of feedback and real-time analysis of game-level performance dynamics contribute to personalized gameplay experiences [55], as shown in Figure 10(b).

VCoach [56] takes a similar approach by analyzing objective user response states such as punching speed, reaction time, and punching posture during exercise. These data were compared to pre-collected data from professional boxing coaches, enabling the system to generate adaptive and

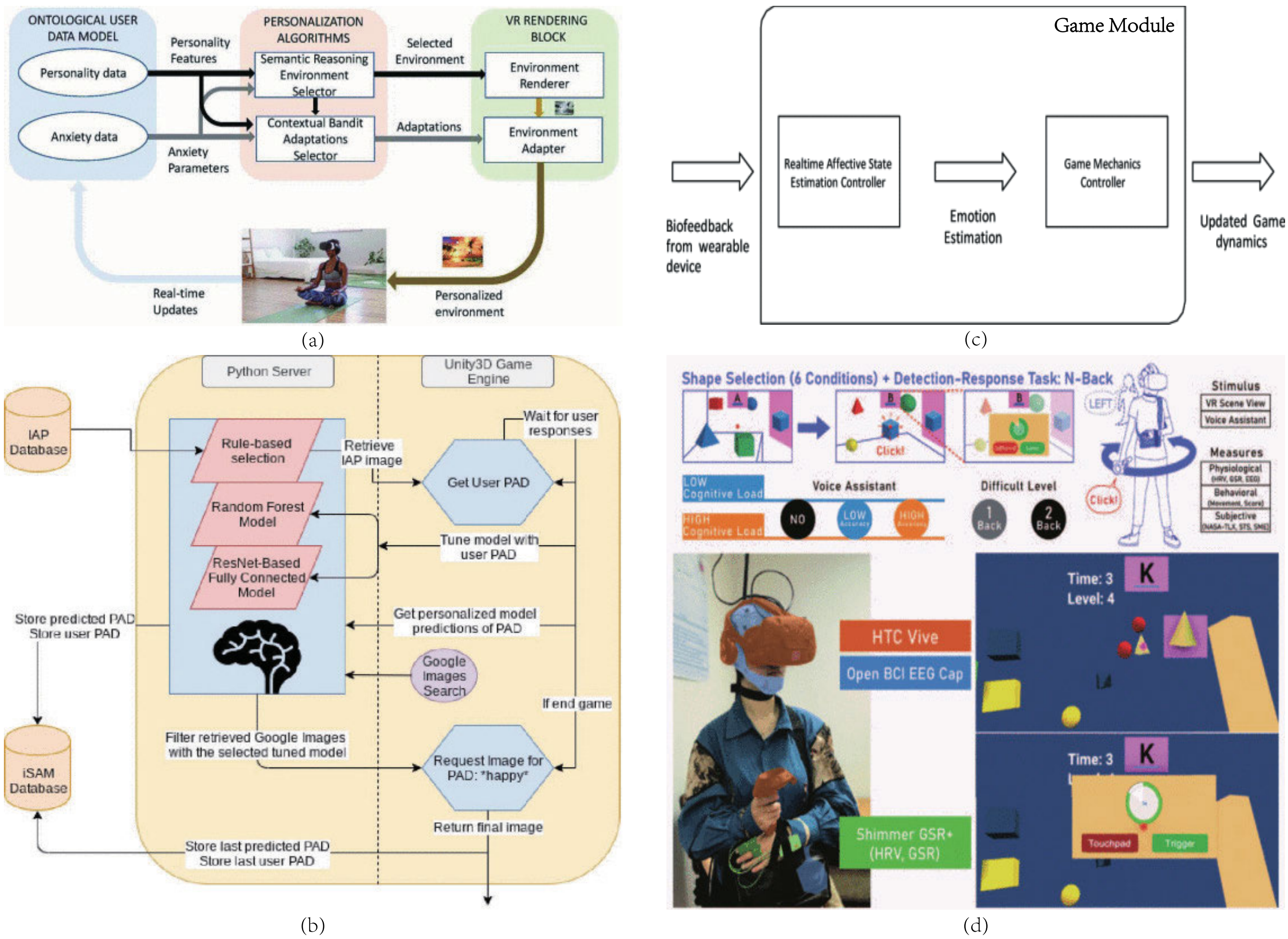


FIGURE 9. Emotional content personalization. (a) Block representation of a custom VR system for relaxing [48]. (b) Logic flow in iSAM [49]. (c) The elements that make up the game module are described [51]. (d) Design, setup, and tasks for experimental research focusing on the psychophysical link between cognitive load and VR's virtual assistants [52].

personalized boxing training routines in real-time, as shown in Figure 10(c).

When using DTBVis [60] technology, experts can better understand the similarities and differences between DTB and the human brain by supporting iterative exploration at various levels and granularities, along with automatic similarity recommendation and high-dimensional exploration. This allows them to customize the model and improve its functionality.

User interaction data in VR supermarkets can be analyzed using techniques such as Neural Collaborative Filtering (NCF) to provide personalized recommendations [57], as shown in Figure 10(d). By leveraging collaborative filtering neural networks, the system can generate personalized product recommendations based on invisible ratings derived from user data.

Privacy considerations are essential in personalized content generation, and techniques like Convolutional Neural Networks (CNNs) can be employed to ensure long-term effectiveness while minimizing the amount of user data required for training [58], as shown in Figure 10(e).

By harnessing objective data from user interactions, VR systems can deliver highly personalized content that caters to individual preferences and needs. However, factors such as the availability and quality of data, privacy concerns, and the challenge of accurately capturing and interpreting user interactions can affect the effectiveness of personalized content delivery. Despite these limitations, leveraging objective data in VR systems can significantly enhance user engagement, immersion, and satisfaction by providing tailored experiences that align with specific requirements and objectives.

V. INTELLIGENT AGENTS

Scientists have developed intelligent agents that focus on content generation to enhance the central and immersive nature of content generation. Intelligent agents enable virtual animation to take on the character of a main storyline, whereas AI integration enhances the simulation, and the adaptive nature of the intelligent agent avoids the frustration of mechanical animation. Intelligent agents can be applied in three main areas: those based on data training, those

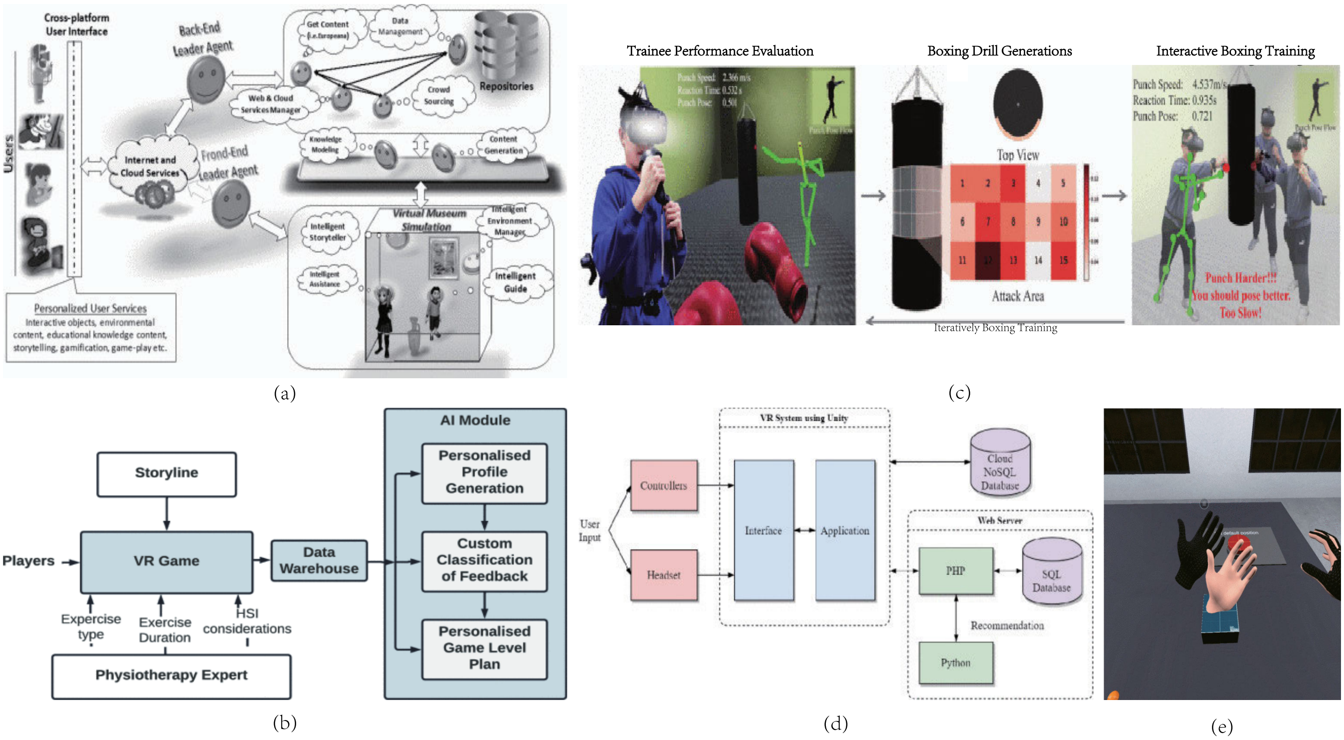


FIGURE 10. Non-emotional content personalization. (a) Ideally suited for dynamic personal virtual worlds based on several intelligences [54]. (b) Use of AI and VR to structurally visualize a suggested framework for individualized physical therapy rehabilitation [55]. (c) Overview of the virtual customized boxing training system VCoach, comprising student performance evaluation, boxing exercise production, and interactive boxing exercise. For iterative skill improvement, trainees can engage in virtual boxing instruction using a consumer-grade wearable VR device, such as the HTC Vive [56]. (d) The VR supermarket's suggested layout [57]. (e) The visualization shows the position and rotation of the genuine hand and the computer-generated hand with vivid skin tones (solid black). The motion being practiced is elbow bending, and the separation between the two hands is determined by varying the rotational angle of the elbow itself [58].

constructed from external knowledge systems, and those trained to adapt to interactive data.

A. INTELLIGENT AGENTS TRAINED BASED ON DATA

In the realm of content generation, the data generated during the development and evolution of VR content serve as a valuable resource for training intelligent agents. Researchers have leveraged the empirical data obtained from VR content to enhance the intelligence of these agents. By allowing the data to evolve in real time, the performance of intelligent agents can progressively improve, resulting in more realistic and sophisticated content generation.

One notable approach is the application of artificial neural networks (ANN) for training intelligent agents. Paladin [61], employed an ANN to control the agents' behavior. The predicted positions obtained from the neural networks are evaluated for their reasonableness, allowing intelligent agents to adjust their actions accordingly. This adaptive behavior, combined with pre-programmed behaviors, enhances the overall intelligence of the agents, as illustrated in Figure 11(a).

Another instance involves the use of multilayer perceptron (MLP) neural networks in mastering fundamental tennis-playing abilities [62]. By collecting essential information from a simulated environment, intelligent agents are trained using MLP-based neural networks, enabling them to acquire the necessary skills to play tennis proficiently, as shown in Figure 11(b).

Furthermore, an RBF_LA learning algorithm was proposed for tennis matches [63]. This algorithm incorporates different learning strategies under various conditions, including no learning process, learning after a match, and learning during a match. By adapting their strategies based on real-time feedback, intelligent agents can improve their performance in tennis matches, as depicted in Figure 11(c).

In the context of virtual-scene navigation, agents are trained to learn and navigate diverse paths [64], as shown in Figure 11(d). Through a reward-based training method, real-time navigation systems enable agents to comprehend and adapt to obstacles, path safety, traffic conditions, and other relevant data, resulting in efficient and effective pathfinding.

Augmented learning techniques have also been employed to train intelligent agents. For instance, Linqin Cai et al. successfully trained agents to navigate through a maze using reinforcement learning (RL) and a Pixy camera sensor to detect and interact with different objects in the environment [65]. This approach grants the agents the freedom to choose optimal routes, fostering adaptability and intelligence in their decision-making processes.

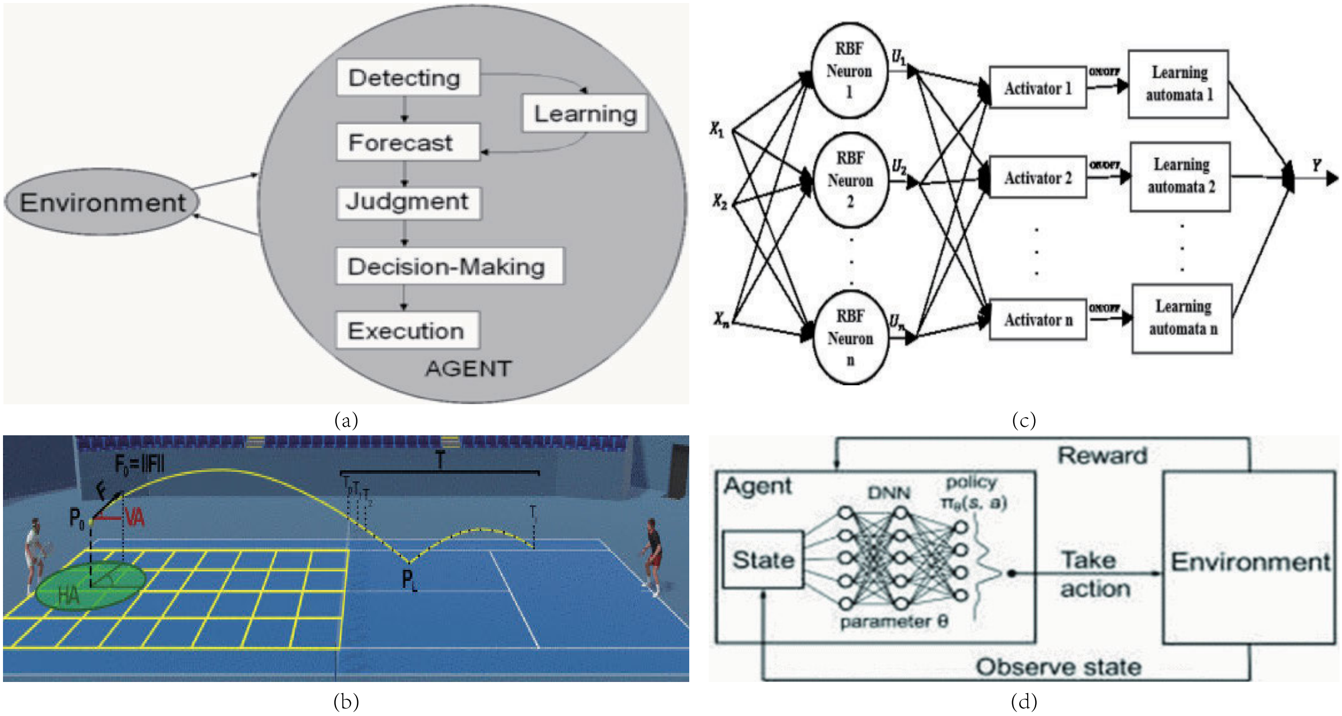


FIGURE 11. Intelligent agents trained based on data. (a) Program for the agency [61]. (b) Building data sets on a virtual tennis court [62]. (c) Structure of the RBF_LA suggested model [63]. (d) Incorporating agents into virtual worlds [64].

Content generation in VR environments can be significantly enhanced by utilizing data-driven training methodologies and intelligent agents. However, the integration of real-time data and adaptive mechanisms may introduce complexities in terms of computational resources and system responsiveness. Despite these limitations, the utilization of data-driven training methodologies and intelligent agents holds great potential for improving content generation and creating more immersive experiences in metaverse environments.

B. INTELLIGENT AGENTS CONSTRUCTED FROM EXTERNAL KNOWLEDGE SYSTEMS

To ensure compatibility between intelligent agents and the logic of reality, external knowledge systems can be leveraged to enhance an agent's understanding of logic. By integrating a real-world logic system, VR simulations can be more closely aligned with real-world principles. This approach can be realized through the construction of intelligent agents using external knowledge systems.

One such approach is the FuSM (Fuzzy State Machine) [66]. By simulating emotions and controlling the behavior of characters affected by these emotions, the FuSM approach creates high-performance AI emotions and AI systems, as illustrated in Figure 12(a). This technology allows for the simultaneous of multiple emotions, adding an element of unpredictability that enhances the credibility of virtual characters and robots. FuSM effectively addresses the combinatorial explosion of DFSM (Deterministic Finite State Machine) states while offering significant expressiveness.

In the case of the Intelligent Virtual Human Animation System (IVHAS) [67], which is based on the Semantic Web, virtual scenes are defined with rich semantic information. This enables computers to comprehend the significance of virtual scenes and perform automated analytical processing, as depicted in Figure 12(b). Semantic virtual environments incorporate various aspects of human behavior, perception, and behavioral planning in the real world, providing rich semantic information to objects within virtual scenarios.

Similarly, the intelligent teaching module of the Smart Physics Lab [12] presents its domain knowledge externally to the Intelligent VR Teaching System (IVRTS) [12], as shown in Figure 12(c). Utilizing external domain knowledge, the teaching module enhances its capabilities and provides a more comprehensive and intelligent virtual teaching experience.

By integrating external knowledge systems, intelligent agents in VR environments can align their logic and behavior with real-world principles, resulting in more realistic and contextually relevant interactions. Nevertheless, challenges may arise in effectively integrating and updating external knowledge sources, ensuring the accuracy and relevance of the information utilized by agents, and addressing potential biases or limitations in the knowledge systems themselves. Nonetheless, the integration of external knowledge systems is expected to improve the intelligence and authenticity of intelligent agents in a metaverse environment.

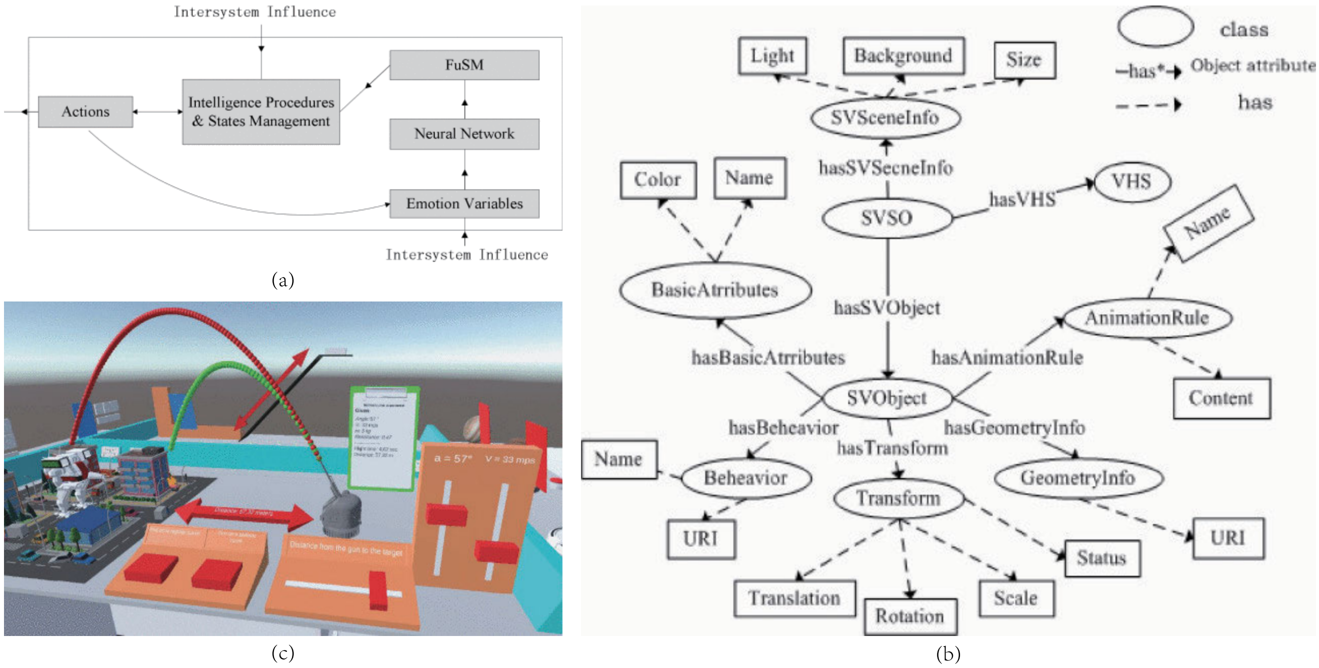


FIGURE 12. Intelligent agents constructed from external knowledge systems. (a) The design of artificial emotion and AI systems [66]. (b) An illustration of the Semantic Virtual Scene Ontology (SVSO) in schematic form [67]. (c) IVRTS interface and a ballistics lab [12].

C. INTELLIGENT AGENTS TRAINED FOR INTERACTION DATA ADAPTATION

In addition to the data generated during the development of VR scene content, real-world data generated by users through their interactions also play a crucial role. Researchers can collect and leverage relevant data to make intelligent agents more contextually relevant to users' situations, thereby enhancing the overall immersive experience. There are three primary situations in which interaction data can be collected to train intelligent agents: simulating realistic images, adapting agent behavior based on interaction actions, and adapting agent collaboration in task processes.

1) SIMULATING REALISTIC IMAGES

To facilitate a more immersive experience, operators can simulate an image of a real person by leveraging the image features of an actual individual. In the case of a virtual companion dining system designed for the elderly [68], researchers have implemented a technique that involves segmenting the RGB images of a real person's area using depth information. By utilizing this depth information, they can restore the point cloud of a person's area and rebuild the grid of virtual images. Furthermore, they mapped the segmented color information onto the virtual body based on texture-mapping rules, as depicted in Figure 13(a). This approach satisfies the emotional need for enhanced realism in the virtual companion, resulting in a more authentic and engaging user experience.

In the context of avatar creation, researchers such as Miao-Chi Liu Chang et al. explored an alternative approach to obtaining pose data without relying on skeletal movement tracking. Instead, they trained a posture classifier using Google's experimental AI API, 'Teachable Machine' [69]. By leveraging user-recognized poses, they enable the interactive creation of avatars in a virtual space, as illustrated in Figure 13(b). This methodology provides improved outcomes and a more efficient user experience.

Simulating real images using real-world data is an effective technique that can enhance immersion in a metaverse environment and meet users' emotional needs to enhance realism. However, there are challenges in obtaining and processing large-scale and diverse real-world data to ensure its accuracy and representativeness. The fidelity of an analog image may be constrained by the quality and resolution of the input data. Nevertheless, incorporating data into image simulation methods has a positive effect.

2) INTERACTION ACTIONS TO ADAPT AGENT CHANGES

To cater to different operator categories, it is important to adapt the agent's approach based on the user's interaction actions. Recognizing this, researchers have devised methods to evaluate users during their interactions and accordingly select the most appropriate agent approach.

In the context of home rehabilitation games, Elor and Kurniawan developed an upper limb exercise component that assists users in learning and guiding their exercise movements [70]. To personalize the rehabilitation strategy and adapt to the exercise difficulty and assistance, they employed a technique called generative imitation learning (GAIL) and proximal strategy optimization (PPO) with the use of virtual butterflies, as depicted in Figure 13(c). By leveraging

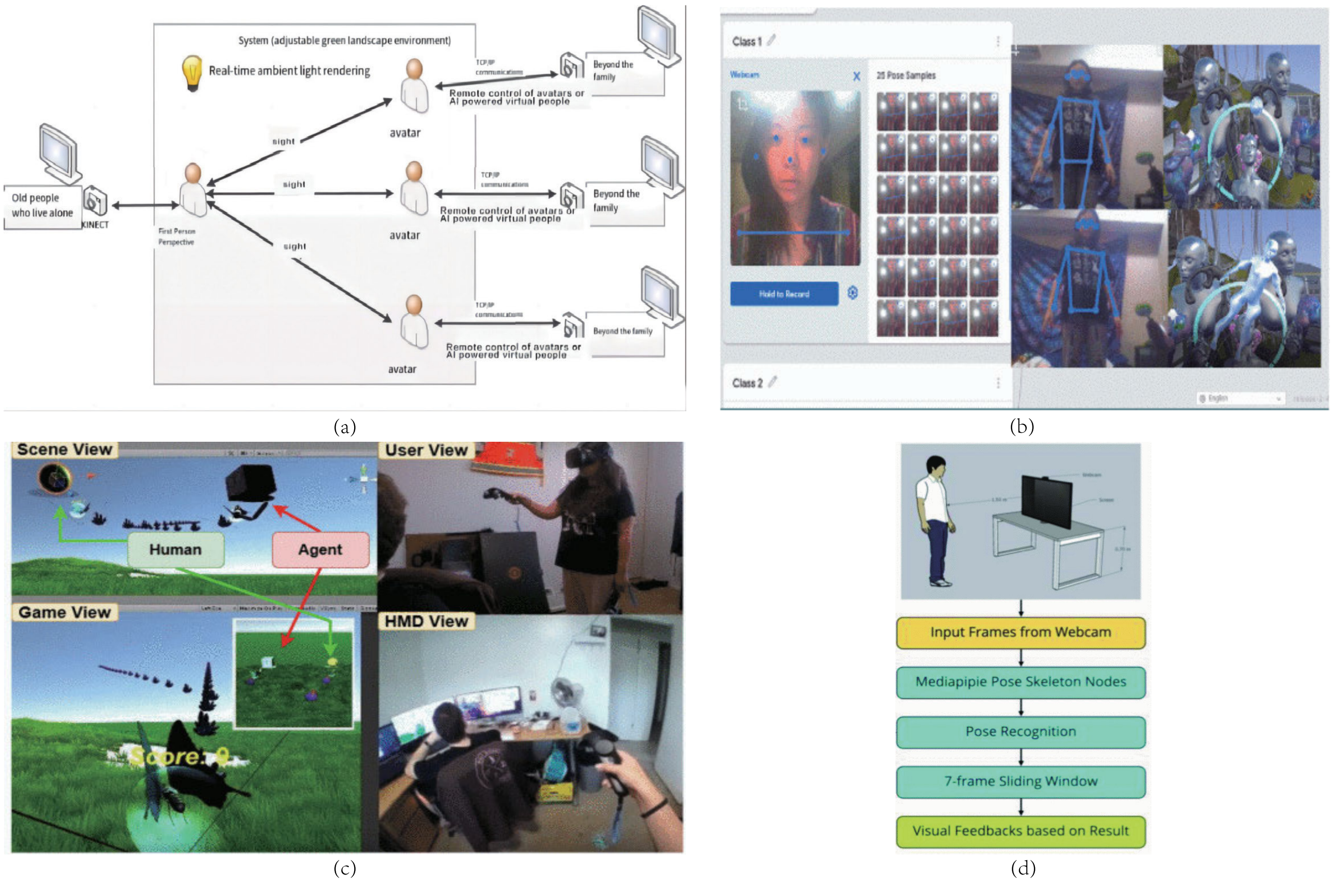


FIGURE 13. Simulating realistic images. (a) The figure of the system architecture for a senior dining system using virtual avatars [68]. (b) Body skeleton joints were found in the recognized produced avatar (right side) as well as in frames taken from the built-in camera (left and center portions) [69]. (c) IB game competition with skilled players. While defending the butterfly, the user competes with the project IB agent to acquire the most crystals. The user's job is to safeguard the butterfly, and the agent is positioned to the user's right. The scene and game views display crystal routes as well as avatars of people and agents [70]. (d) The 'Ant-Man Vision' experience's method [71].

these approaches, exercise difficulty and assistance can be dynamically adjusted based on user performance, ensuring a personalized rehabilitation experience.

In the case of museum anthropomorphic vision [71], researchers took a different approach to process user interactions, as shown in Figure 13(d). They trained a behavioral model using Long Short-Term Memory (LSTM) and utilized cameras and 3D skeleton position acquisition to detect user behavior in real space. Based on this information, they modified the state changes of the intelligent agent, Ant-Man, in the museum. This adaptive approach enables the agent to respond to user actions and to create a more immersive and interactive museum experience.

By incorporating interaction actions to adapt to agent changes, researchers can tailor the agent's behavior and responses to individual users, creating a more personalized and engaging user-agent interaction.

3) INTERACTION PROCESSES TO ADAPT TO TASK COLLABORATION

The interaction processes not only fine-tune the real-time behavior of the agent but also address the issue of task collaboration, adding unpredictability and interest to the VR experience. Various approaches have been employed to address this challenge and to enhance task collaboration in virtual environments.

For instance, in the case of Paladin, the challenge of work distribution for cooperation was addressed by utilizing ESP neural networks [72], [76], as depicted in Figure 14(a). This approach allows efficient and balanced task allocation among cooperative agents, ensuring smooth collaboration.

In the context of the IVTS [73], a declarative form based on Petri Nets (PNs) [77] was utilized to describe training task planning. Researchers have developed an algorithm to create Task Planning Petri Nets (TP-PNets) and established a Hierarchical Coloured Petri Net (HCPN) model to evaluate agent task planning behavior, as shown in Figure 14(b). These measures provided dependability and flexibility in task planning, allowing agents to adapt their collaboration strategies based on the specific requirements of the tasks.

With the increase in the complexity of VR content, researchers have focused on creating sophisticated virtual internal agent models to facilitate intelligent and engaging

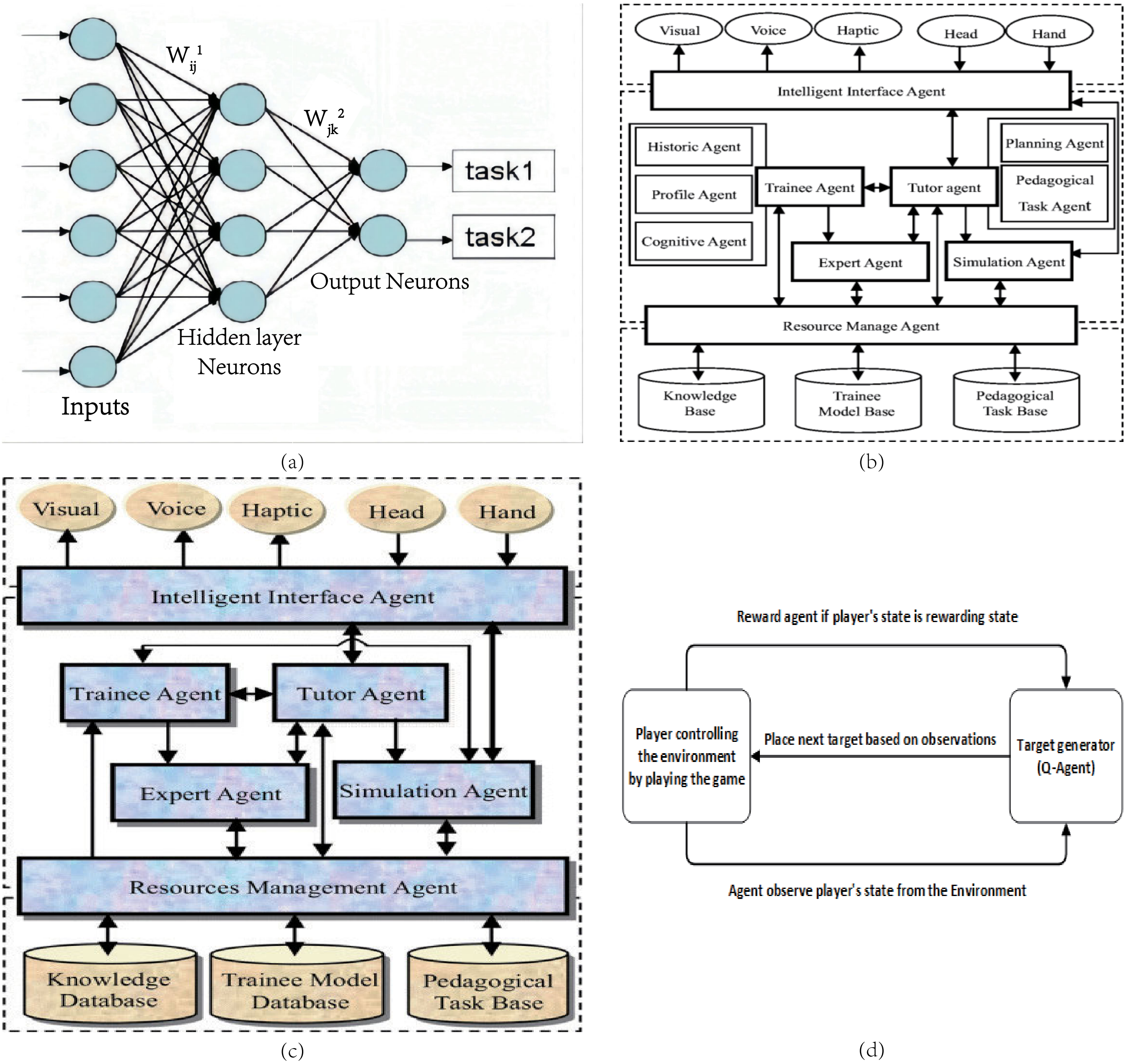


FIGURE 14. Interaction processes to adapt task collaboration. (a) Network collaborations. The artificial neural network uses the values that the agents have received to conduct tasks on the agents by the network's outputs. The weights of the linkages from input neuron I to hidden neuron j are shown by W_{ij}^{1}, while the links from hidden neuron j to output neuron k are indicated by W_{ij}^{2} [72]. (b) The IVTS suggested architectural [73]. (c) An intelligent virtual training system's framework [74]. (d) Using Q-learning, the target location generator architecture [75].

VR systems. A highly competent avatar model was developed in the field of mine safety instruction [74], as illustrated in Figure 14(c). This model incorporates perceptual, behavioral, mental state, and cognitive modules, enabling the avatar to execute training orders, exhibit the desired physical abilities, and mediate interactions between real-world and virtual contexts.

In the Flying with Friends motor disorder rehabilitation system [75], the Deep Q-Networks (DQN) approach was employed to tackle complex tasks, as shown in Figure 14(d). This approach has proven to be an effective choice for improving task collaboration and achieving better rehabilitation outcomes.

While accurately modeling and predicting user intentions and preferences, as well as ensuring seamless coordination and synchronization among multiple agents, can be complex, especially in dynamic and unpredictable environments, researchers aim to enhance the collaborative behavior of intelligent agents in virtual environments by incorporating interaction processes. This approach aims to create a dynamic, engaging, and realistic user experience. Despite these challenges, there is optimism that the combination of interaction processes and task collaboration can lead to significant improvements in the collaborative behavior of intelligent agents, offering users a more immersive and satisfying metaverse environment.

VI. DISCUSSION AND CONCLUSION

This paper introduced innovative aspects in the field of intelligent visual content generation, providing a comprehensive overview of AI methods in scene content generation, simulated biology, personalized content, and intelligent agents. The integration of AI algorithms with real-world data and external knowledge systems is highlighted, enabling the creation of authentic and contextually rich virtual scenarios. This study emphasizes the importance of interdisciplinary collaboration and considers human cognitive and emotional factors. Key limitations and challenges were identified, inspiring future research on creative content generation, ethics, and computational efficiency. Overall, this study enriches the knowledge and drives the development of immersive visual content in virtual environments.

Intelligent visual-content generation has significantly impacted content creation and user experience by enabling the creation of visually stunning scenes, personalized content, and adaptive intelligent agents in virtual environments. In the following sections, we summarize and discuss three potential research trends in the construction of intelligent metaverse scene content, namely the improvement of personalized methods and the integration of brain-computer interfaces, the integration of AI with real-world data and external knowledge systems, and the importance of interdisciplinary collaboration.

A. POTENTIAL TREND DISCUSSION: IMPROVEMENT OF PERSONALIZED METHODS AND INTEGRATION OF BRAIN-COMPUTER INTERFACES

The demand for personalized content tailored to individual user preferences and emotions is increasing. With the continuous improvement AI algorithms for recognizing human emotions, methods such as speech emotion recognition [78], physiological signal emotion recognition [79], facial image emotion recognition [80], text mixing analysis [81], cultural subdivision and overall analysis [82] have significantly enhanced the accuracy of emotion recognition. The advancement of formulaic feedback theory [83], [84] has also contributed to more perfect adaptive emotion generation in metaverse scenes. Additionally, the rapid development of brain-computer interface technology [85] provides the possibility for real-time emotional feedback, ensuring the real-time and accuracy of adaptive changes in scenes. Integrating intelligent analysis of electroencephalogram (EEG) data [86] can further enhance the efficiency and diversity of content generation, resulting in more realistic and personalized metaverse content.

B. POTENTIAL TREND DISCUSSION: INTEGRATION OF AI WITH REAL-WORLD DATA AND EXTERNAL KNOWLEDGE SYSTEMS

In the future, a deeper analysis of data relationships [87], more refined data collection and simulation [88], extraction and growth of anthropomorphic knowledge [89], and the cross-fusion analysis of knowledge data obtained from diverse real-world sources [90] will provide new possibilities for the intelligent generation of three-dimensional simulation metaverse scenes. The integration of AI algorithms with real-world data and external information systems will continue to improve, enabling the creation of more realistic and contextually rich virtual scenes. This integration will lead to the development of smarter and more creative agents, thereby enhancing the overall immersive experience for users.

C. POTENTIAL TREND DISCUSSION: INTERDISCIPLINARY COLLABORATION

Collaboration between different disciplines, such as psychology, neuroscience, and computer science, will play a crucial role in enhancing our understanding of the cognitive and emotional impact of intelligent metaverse scene content. The demand for intelligent visual content generation in scientific fields such as education [91], medicine [92], chemistry [93], physics [94], and geology [95] is expected to increase. Intelligent content generation can help beginners in various disciplines gain a fast, effective, and interesting understanding, while providing experts with a more intuitive, rich, and immersive experience. In humanities fields, such as education and art, content sensitivity outweighs accuracy. For example, cross-modal metaverse scene generation, including direct text generation [96], will find popularity in artistic fields like ancient poetry and artistic styles. In the scientific and engineering fields, accuracy takes precedence over receptivity. For instance, intelligent construction of biologically variable metaverse models based on medical knowledge services [97] may become a direction of interest.

In conclusion, the generation of intelligent metaverse scene content has transformed content creation and the user experience. Continued advancements in AI algorithms and interdisciplinary collaborations hold promise for further improvement. This survey aims to stimulate more ideas about scene content construction through the use of intelligent metaverse immersive technology.

REFERENCES

[1] X. Li, H. Duan, B. Liu, X. Wang, and F.-Y. Wang, "A novel framework to generate synthetic video for foreground detection in highway surveillance scenarios," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 6, pp. 5958–5970, Jun. 2023.

[2] Y. Zhao, J. Jiang, Y. Chen, R. Liu, Y. Yang, X. Xue, and S. Chen, "Metaverse: Perspectives from graphics, interactions and visualization," *Vis. Informat.*, vol. 6, no. 1, pp. 56–67, Mar. 2022.

[3] X. Li and Y. Shi, "Computer vision imaging based on artificial intelligence," in *Proc. Int. Conf. Virtual Reality Intell. Syst. (ICVRIS)*, Aug. 2018, pp. 22–25.

[4] R. Liu, M. Gao, L. Wang, X. Wang, Y. Xiang, A. Zhang, J. Xia, Y. Chen, and S. Chen, "Interactive extended reality techniques in information visualization," *IEEE Trans. Hum.-Mach. Syst.*, vol. 52, no. 6, pp. 1338–1351, Dec. 2022.

[5] L. Li, "The impact of artificial intelligence painting on contemporary art from disco diffusion's painting creation experiment," in *Proc. Int. Conf. Frontiers Artif. Intell. Mach. Learn. (FAIML)*, Jun. 2022, pp. 52–56.

[6] X. Shi and W. Hu, "A survey of image style transfer research," in *Proc. 2nd Int. Conf. Comput. Sci., Electron. Inf. Eng. Intell. Control Technol. (CEI)*, Sep. 2022, pp. 133–137.

[7] J. P. A. Campos and R. Rieder, "Procedural content generation using artificial intelligence for unique virtual reality game experiences," in *Proc. 21st Symp. Virtual Augmented Reality (SVR)*, Oct. 2019, pp. 147–151.

[8] C. Tang, Z. Wang, X. Sima, and L. Zhang, "Research on artificial intelligence algorithm and its application in games," in *Proc. 2nd Int. Conf. Artif. Intell. Adv. Manuf. (AIAM)*, Oct. 2020, pp. 386–389.

[9] Y. Zhang, G. Zhang, and X. Huang, "A survey of procedural content generation for games," in *Proc. Int. Conf. Culture-Oriented Sci. Technol. (CoST)*, Aug. 2022, pp. 186–190.

[10] F. Lo, F. Su, S. Chen, J. Qiu, and J. Du, "Artificial intelligence aided innovation education based on multiple intelligence," in *Proc. IEEE Int. Conf. Artif. Intell., Robot., Commun. (ICAIRC)*, Jun. 2021, pp. 12–15.

[11] K. Kroes, I. Saccardi, and J. Masthoff, "Empathizing with virtual agents: The effect of personification and general empathic tendencies," in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Dec. 2022, pp. 73–81.

[12] F. Dudyrev, O. Maksimenkova, and D. Mikhailenko, "Intelligent virtual reality tutoring systems as a new generation of simulators: Requirements and opportunities," in *Proc. IEEE Global Eng. Educ. Conf. (EDUCON)*, Apr. 2021, pp. 706–718.

[13] R. Kaviyaraj and M. Uma, "A survey on future of augmented reality with AI in education," in *Proc. Int. Conf. Artif. Intell. Smart Syst. (ICAIS)*, Mar. 2021, pp. 47–52.

[14] R. Wang, L. Jiang, J. Yuan, X. Xu, and W. Wang, "Virtual reality scene construction based on multimodal video scene segmentation algorithm," in *Proc. IEEE 8th Joint Int. Inf. Technol. Artif. Intell. Conf. (ITAIC)*, May 2019, pp. 1817–1820.

[15] A. Becher, C. Axenie, and T. Grauschopf, "VIRTOOAIR: Virtual reality TOOlbox for avatar intelligent reconstruction," in *Proc. IEEE Int. Symp. Mixed Augmented Reality Adjunct (ISMAR-Adjunct)*, Oct. 2018, pp. 275–279.

[16] Y. Zhang, "Application of interactive artificial intelligence in data evaluation of ink animation production," in *Proc. Int. Conf. Appl. Artif. Intell. Comput. (ICAAIC)*, May 2022, pp. 34–37.

[17] T. Fréville, C. Hamesse, B. Pairet, R. Lahouli, and R. Haelterman, "From floor plans to virtual reality," in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Nov. 2021, pp. 129–133.

[18] N. Matsuo and M. Imura, "Turning a messy room into a fully immersive VR playground," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Abstr. Workshops (VRW)*, Mar. 2021, pp. 759–760.

[19] S. Naritomi, R. Tanno, T. Ege, and K. Yanai, "FoodChangeLens: CNN-based food transformation on HoloLens," in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Dec. 2018, pp. 197–199.

[20] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2242–2251.

[21] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham, Switzerland: Springer, 2015, pp. 234–241.

[22] D. Holden, T. Komura, and J. Saito, "Phase-functioned neural networks for character control," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1–13, Aug. 2017.

[23] W. Xu, I. Alarab, C. Lloyd-Buckingham, S. Bowden, B. Noer, F. Charles, S. Prakoonwit, A. Callaway, S. Ellis, and C. Jones, "Re-enacting football matches in VR using virtual agents' realistic behaviours," in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Dec. 2022, pp. 119–123.

[24] R. Guo, J. Cui, W. Zhao, S. Li, and A. Hao, "Hand-by-hand mentor: An AR based training system for piano performance," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Abstr. Workshops (VRW)*, Mar. 2021, pp. 436–437.

[25] J. Lugrin, M. Cavazza, M. Palmer, and S. Crooks, "Artificial intelligence-mediated interaction in virtual reality art," *IEEE Intell. Syst.*, vol. 21, no. 5, pp. 54–62, Sep. 2006.

[26] T. Palazzolo, A. Lemke, S. Saad, C. Zhang, S. P. Jayanti, J. O'Shea, and R. G. Reynolds, "Deepdive: Using AI, machine learning, and virtual reality to explore ancient submerged civilizations," in *Proc. 3rd Int. Conf. Artif. Intell. Industries*, Sep. 2020, pp. 79–82.

[27] M. I. Iglesias, M. Jenkins, and G. Morison, "An enhanced photorealistic immersive system using augmented situated visualization within virtual reality," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Abstr. Workshops (VRW)*, Mar. 2021, pp. 514–515.

[28] F. Silva, M. A. Togou, and G.-M. Muntean, "An innovative algorithm for improved quality multipath delivery of virtual reality content," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB)*, Oct. 2020, pp. 1–6.

[29] V. Franchi and E. Ntagiou, "Augmentation of a virtual reality environment using generative adversarial networks," in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Nov. 2021, pp. 219–223.

[30] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," in *Proc. ICML*, 2015, vol. 2, no. 1, pp. 1–30.

[31] F. Gao and G. Xu, "Research on multi-mode interactive virtual reality scene enhancement based on artificial intelligence," in *Proc. 12th Int. Conf. Measuring Technol. Mechatronics Autom. (ICMTMA)*, Feb. 2020, pp. 391–396.

[32] L. Yang, Y.-C. Tse, P. V. Sander, J. Lawrence, D. Nehab, H. Hoppe, and C. L. Wilkins, "Image-based bidirectional scene reprojection," *ACM Trans. Graph.*, vol. 30, no. 6, pp. 1–10, Dec. 2011, doi: 10.1145/2070781.2024184.

[33] A. Pilco, X. Luo, A. A. N. Newball, C. Zuniga, and C. Lozano-Garzón, "Procedural animation generation technology of virtual fish flock," in *Proc. Int. Conf. Virtual Reality Visualizat. (ICVRV)*, Nov. 2019, pp. 233–237.

[34] D. Cracknell, M. P. White, S. Pahl, W. J. Nichols, and M. H. Depledge, "Marine biota and psychological well-being: A preliminary examination of dose–response effects in an aquarium setting," *Environ. Behav.*, vol. 48, no. 10, pp. 1242–1269, Dec. 2016, doi: 10.1177/0013916515597512.

[35] M. Xian-Yu, B. Xiao-Juan, Z. Shu-Jun, and T. Xu-Yan, "Design of multi-perception system model for artificial fish in virtual environment," in *Proc. IEEE Int. Conf. Ind. Informat.*, Aug. 2006, pp. 1108–1112.

[36] Q. Xu and M. Zhou, "The modeling of fish roe based on artificial life method," in *Proc. Int. Conf. Mech. Sci., Electric Eng. Comput. (MEC)*, Aug. 2011, pp. 880–882.

[37] Y. Liu, "Research on the cognitive model artificial fish based on virtual marine environment," in *Proc. 8th Int. Conf. Inf. Technol. Med. Educ. (ITME)*, Dec. 2016, pp. 719–723.

[38] J. K. Parrish, S. V. Viscido, and D. Grunbaum, "Self-organized fish schools: An examination of emergent properties," *Biol. Bull.*, vol. 202, no. 3, pp. 296–305, Jun. 2002.

[39] M. H. Mozaffari, S. Wen, and W.-S. Lee, "Virtual reality and tracking the mating behavior of fruit flies: A machine learning approach," in *Proc. IEEE Int. Conf. Comput. Intell. Virtual Environments Meas. Syst. Appl. (CIVEMSA)*, Jun. 2018, pp. 1–5.

[40] T. Lints, "FlockHeadz: Virtual flock in a room used as a controller," in *Proc. IEEE Swarm Intell. Symp.*, Sep. 2008, pp. 1–5.

[41] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," *ACM SIGGRAPH Comput. Graph.*, vol. 21, no. 4, pp. 25–34, Aug. 1987.

[42] H. X. Trejo, "Creating an immersed sheep and wool VR/AI experience," in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Dec. 2019, pp. 237–2371.

[43] G. Noronha, G. Batista, and L. P. Soares, "Crowd simulation with augmented reality," in *Proc. XV Symp. Virtual Augmented Reality*, May 2013, pp. 228–231.

[44] I. Aoki, "A simulation study on the schooling mechanism in fish," *Nippon Suisan Gakkaishi*, vol. 48, no. 8, pp. 1081–1088, 1982.

[45] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *J. Artif. Intell. Res.*, vol. 4, no. 1, pp. 237–285, Jan. 1996.

[46] K. Morihiro, T. Isokawa, H. Nishimura, and N. Matsui, "Characteristics of flocking behavior model by reinforcement learning scheme," in *Proc. SICE-ICASE Int. Joint Conf.*, 2006, pp. 4551–4556.

[47] R. Wangananont, N. Buppodom, S. Chanthanuraks, and V. Kotrajaras, "Simulation of homogenous fish schools in the presence of food and predators using reinforcement learning," in *Proc. 17th Int. Joint Symp. Artif. Intell. Natural Lang. Process. (iSAI-NLP)*, Nov. 2022, pp. 1–6.

[48] J. Heyse, T. D. Jonge, M. T. Vega, F. D. Backere, and F. De Turck, "A personalized virtual reality experience for relaxation therapy," in *Proc. 11th Int. Conf. Quality Multimedia Exper. (QoMEX)*, Jun. 2019, pp. 1–3.
[49] A. Elor and A. Song, "ISAM: Personalizing an artificial intelligence model for emotion with pleasure-arousal-dominance in immersive virtual reality," in *Proc. 15th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, Nov. 2020, pp. 572–576.
[50] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Current Psychol.*, vol. 14, no. 4, pp. 261–292, Dec. 1996.
[51] A. Dingli and L. Bondin, "Realtime adaptive virtual reality for pain reduction," in *Proc. IEEE Conf. Games (CoG)*, Aug. 2019, pp. 1–4.
[52] K. Gupta, "[DC] exploration of context and physiological cues for personalized emotion-adaptive virtual reality," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Abstr. Workshops (VRW)*, Mar. 2022, pp. 960–961.
[53] P. Dhope and M. B. Neelagar, "Real-time emotion recognition from facial expressions using artificial intelligence," in *Proc. 2nd Int. Conf. Artif. Intell. Signal Process. (AISP)*, Feb. 2022, pp. 1–6.
[54] C. Kiourt, G. Pavlidis, A. Koutsoudis, and D. Kalles, "Multi-agents based virtual environments for cultural heritage," in *Proc. 26th Int. Conf. Inf., Commun. Autom. Technol. (ICAT)*, Oct. 2017, pp. 1–6.
[55] T. Kempitiya, D. De Silva, E. Rio, R. Skarbez, and D. Alahakoon, "Personalised physiotherapy rehabilitation using artificial intelligence and virtual reality gaming," in *Proc. 15th Int. Conf. Human Syst. Interact. (HSI)*, Jul. 2022, pp. 1–6.
[56] H. Chen, Y. Wang, and W. Liang, "VCoach: Enabling personalized boxing training in virtual reality," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Abstr. Workshops (VRW)*, Mar. 2022, pp. 546–547.
[57] D. Shravani, Y. R. Prajwal, P. V. Atreyas, and G. Shobha, "VR supermarket: A virtual reality online shopping platform with a dynamic recommendation system," in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Nov. 2021, pp. 119–123.
[58] A. M. Soccini and F. Cena, "The ethics of rehabilitation in virtual reality: The role of self-avatars and deep learning," in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Nov. 2021, pp. 324–328.
[59] E. Osipov, S. Kahawala, D. Haputhanthri, T. Kempitiya, D. D. Silva, D. Alahakoon, and D. Kleyko, "Hyperseed: Unsupervised learning with vector symbolic architectures," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Nov. 16, 2022, doi: 10.1109/TNNLS.2022.3211274.
[60] Y. Li, X. Li, S. Shen, L. Zeng, R. Liu, Q. Zheng, J. Feng, and S. Chen, "DTBVis: An interactive visual comparison system for digital twin brain and human brain," *Vis. Informat.*, vol. 7, no. 2, pp. 41–53, Jun. 2023.
[61] J. Li and Z. Miao, "Agent based intelligent virtual environment and its entertainment application," in *Proc. 8th Int. Conf. Signal Process.*, vol. 2, 2006.
[62] S. M. A. Moazzen and A. Ahmadi, "3D simulation of tennis game using learning algorithms by application of virtual reality," in *Proc. Iranian Conf. Electr. Eng. (ICEE)*, May 2017, pp. 1527–1531.
[63] E. Rahimi and A. Ahmadi, "An AI-based tennis game by application of virtual reality components," in *Proc. Iranian Conf. Electr. Eng. (ICEE)*, May 2017, pp. 2165–2170.
[64] S. Dawn, U. Saraogi, and U. S. Thakur, "Agent-based learning for auto-navigation within the virtual city," in *Proc. Int. Conf. Comput. Perform. Eval. (ComPE)*, Jul. 2020, pp. 007–012.
[65] V. V. R. M. K. Rao, N. Adhikari, and A. D. Ghimire, "Towards training an agent in augmented reality world with reinforcement learning," in *Proc. 17th Int. Conf. Control, Autom. Syst. (ICCAS)*, Oct. 2017, pp. 1884–1888.
[66] J. Li, Z. Wang, and Y. Zhang, "An implementation of artificial emotion based on fuzzy state machine," in *Proc. 3rd Int. Conf. Intell. Human-Mach. Syst. Cybern.*, 2011, pp. 83–86.
[67] C.-T. Song and H. Wang, "Research and implementation of virtual human animation based on distributed semantic environment," in *Proc. IEEE Int. Conf. Electron. Inf. Commun. Technol. (ICEICT)*, Aug. 2016, pp. 78–81.
[68] R. Dai and Z. Pan, "A virtual companion empty-nest elderly dining system based on virtual avatars," in *Proc. IEEE 7th Int. Conf. Virtual Reality (ICVR)*, May 2021, pp. 446–451.
[69] M. L. Chang, Y.-H. Huang, W.-C. Lin, and S.-W. Sun, "Digital fabrication: Machine learning-based immersive experiencing for the virtual space in a future museum," in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Nov. 2021, pp. 102–105.
[70] A. Elor and S. Kurniawan, "Deep reinforcement learning in immersive virtual reality exergame for agent movement guidance," in *Proc. IEEE 8th Int. Conf. Serious Games Appl. Health (SeGAH)*, Aug. 2020, pp. 1–7.
[71] Y.-L. Chen, Y.-X. Cheng, I.-Y. Hung, and S.-W. Sun, "Ant-man vision in the museum with interactive and immersive surreal experience based on machine learning," in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Dec. 2022, pp. 242–244.
[72] L. Jia and M. Zhenjiang, "Entertainment oriented intelligent virtual environment with agent and neural networks," in *Proc. IEEE Int. Workshop Haptic, Audio Vis. Environments Games*, Oct. 2007, pp. 90–95.
[73] L. Cai, T. Mei, Y. Sun, L. Sun, and Z. Ma, "Modeling and analyzing multi-agent task plans for intelligent virtual training system using Petri Nets," in *Proc. 6th World Congr. Intell. Control Autom.*, 2006, pp. 4766–4770.
[74] L. Cai, T. Mei, and Y. Sun, "Multi-agent based intelligent virtual training system," in *Proc. 7th World Congr. Intell. Control Autom.*, 2008, pp. 4226–4230.
[75] D. Hornak, M. Jascur, N. Ferencik, and M. Bundzel, "Proof of concept: Using reinforcement learning agent as an adversary in serious games," in *Proc. IEEE Int. Work Conf. Bioinspired Intell. (IWOBI)*, Jul. 2019, pp. 000111–000116.
[76] B. D. Bryant and R. Miikkulainen, "Neuroevolution for adaotive teams," in *Proc. Congr. Evol. Comput.*, 2003, pp. 2194–2201.
[77] T. Murata, "Petri Nets: Properties, analysis and applications," *Proc. IEEE*, vol. 77, no. 4, pp. 541–580, Apr. 1989.
[78] S.-L. Yeh, Y.-S. Lin, and C.-C. Lee, "A dialogical emotion decoder for speech emotion recognition in spoken dialog," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 6479–6483.
[79] C. Joesph, A. Rajeswari, B. Premalatha, and C. Balapriya, "Implementation of physiological signal based emotion recognition algorithm," in *Proc. IEEE 36th Int. Conf. Data Eng. (ICDE)*, Apr. 2020, pp. 2075–2079.
[80] X. Li and R. Lian, "Correlation modeling of moral emotion based on facial image emotion recognition algorithm," in *Proc. Int. Conf. Augmented Intell. Sustain. Syst. (ICAISS)*, Nov. 2022, pp. 560–562.
[81] M. A. Mahima, N. C. Patel, S. Ravichandran, N. Aishwarya, and S. Maradithaya, "A text-based hybrid approach for multiple emotion detection using contextual and semantic analysis," in *Proc. Int. Conf. Innov. Comput., Intell. Commun. Smart Electr. Syst. (ICSES)*, Sep. 2021, pp. 1–6.
[82] Z. Zhang, L. Xie, and Z. Zhao, "Temporal integration of emotion perception for cross-cultural and multi-emotion music," in *Proc. Int. Conf. Culture-Oriented Sci. Technol. (ICCST)*, Nov. 2021, pp. 198–203.
[83] A. Dixit, M. Vashishtha, and K. Guleri, "An AI based formulated feedback-system for interpreting conclusive emotions for a group of people," in *Proc. IEEE Global Conf. Comput., Power Commun. Technol. (GlobConPT)*, Sep. 2022, pp. 1–5.
[84] T. Akahori, K. Dohsaka, M. Ishii, and H. Ito, "Efficient creation of Japanese tweet emotion dataset using sentence-final expressions," in *Proc. IEEE 3rd Global Conf. Life Sci. Technol. (LifeTech)*, Mar. 2021, pp. 501–505.
[85] A. S. M. S. Reaj, M. Maniruzzaman, and A. A. J. Jim, "Emotion recognition using EEG-based brain computer interface," in *Proc. Int. Conf. Electron., Commun. Inf. Technol. (ICECIT)*, Sep. 2021, pp. 1–4.
[86] J. Li, S. Qiu, Y.-Y. Shen, C.-L. Liu, and H. He, "Multisource transfer learning for cross-subject EEG emotion recognition," *IEEE Trans. Cybern.*, vol. 50, no. 7, pp. 3281–3293, Jul. 2020.
[87] K. Liu, M. Yang, X. Li, K. Zhang, X. Xia, and H. Yan, "M-data-fabric: A data fabric system based on metadata," in *Proc. IEEE 5th Int. Conf. Big Data Artif. Intell. (BDAI)*, Jul. 2022, pp. 57–62.
[88] X. Zhou, "Tennis motion precision video monitoring simulation based on artificial intelligence algorithm," in *Proc. Int. Conf. Artif. Intell. Auto. Robot Syst. (AIARS)*, Jul. 2022, pp. 211–215.
[89] A. S. Ahmad, "Brain inspired cognitive artificial intelligence for knowledge extraction and intelligent instrumentation system," in *Proc. Int. Symp. Electron. Smart Devices (ISESD)*, Oct. 2017, pp. 352–356.
[90] L. M. Alkwai, "An explainable artificial-intelligence-based CNN model for knowledge extraction from the social Internet of Things: Proposing a new model," *IEEE Syst., Man, Cybern. Mag.*, vol. 8, no. 4, pp. 48–51, Oct. 2022.
[91] Z. Dong, "Roman research progress and prospects of 'artificial intelligence + education' area based on CITESPACE knowledge spectrum," in *Proc. 3rd Int. Conf. Educ., Knowl. Inf. Manag.*, 2022, pp. 331–334.
[92] X. Su, Z. You, D. Huang, L. Wang, L. Wong, B. Ji, and B. Zhao, "Biomedical knowledge graph embedding with capsule network for multi-label drug-drug interaction prediction," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 6, pp. 5640–5651, Jun. 2023.

[93] X. Zhao, J. Greenberg, S. McClellan, Y.-J. Hu, S. Lopez, S. K. Saikin, X. Hu, and Y. An, ''Knowledge graph-empowered materials discovery,'' in *Proc. IEEE Int. Conf. Big Data*, Dec. 2021, pp. 4628–4632.

[94] T. Ma, L. Miao, Y. Du, and H. Wang, ''Calibration data prediction of wind tunnel balance based on regression tree,'' in *Proc. 16th Int. Conf. Intell. Syst. Knowl. Eng. (ISKE)*, Nov. 2021, pp. 362–367.

[95] Q. Wang, A. P. Doulgeris, and T. Eltoft, ''Physics-aware training data to improve machine learning for sea ice classification from Sentinel-1 SAR scenes,'' in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2022, pp. 4992–4995.

[96] S. Zhang, J. Yuan, M. Liao, and L. Zhang, ''text2video: Text-driven talking-head video synthesis with personalized phoneme—Pose dictionary,'' in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2022, pp. 2659–2663.

[97] M. Zhang, Z. Yang, C. Liu, and L. Fang, ''Traditional Chinese medicine knowledge service based on semi-supervised BERT-BiLSTM-CRF model,'' in *Proc. Int. Conf. Service Sci. (ICSS)*, Aug. 2020, pp. 64–69.

JUNXIANG WANG is currently pursuing the bachelor's degree with Nanjing Normal University. He has been the author of several software copyrights. His current research interests include artificial intelligence, visualization, virtual reality, deep learning, and content generation.

SIRU CHEN is currently pursuing the degree with Nanjing Normal University. Her current research interests include scientific data visualization and immersive visualization. She is the cofourth author of the 2021 IEEE Pacific VIS Best Poster paper. She is also the author of the paper Multi-User Collaborative Volume Data Illustration and Visualization presented at IEEE Visualization 2020 (IEEE VIS).

YUXUAN LIU is currently pursuing the bachelor's degree with Nanjing Normal University. Her current research interests include visualization and virtual reality.

RICHEN LAU received the Ph.D. degree from Peking University, in 2017. He is an Associate Professor with Nanjing Normal University. He has published more than 30 peer-reviewed journals and conference papers, including ACM CHI, IEEE VIS, IEEE ISMAR, IEEE PacificVis, IEEE THMS, and Bioinformatics. He was a PC/TPC Member of IEEE VIS (short papers), ACHI, ChinaVis, and HPC China. He has supervised students to receive the Best Survey Paper Award from ChinaVis 2019 and the Best Poster Award from IEEE PacificVis.