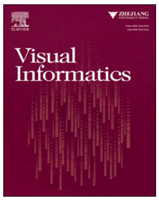
Visual Informatics 6 (2022) 56–67

Contents lists available at ScienceDirect

Visual Informatics

journal homepage: www.elsevier.com/locate/visinf



Metaverse: Perspectives from graphics, interactions and visualization

Yuheng Zhao [a], Jinjing Jiang [a], Yi Chen [b], Richen Liu [c], Yalong Yang [d], Xiangyang Xue [a], Siming Chen [a],*

[a] *School of Data Science, Fudan University, China*
[b] *Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing Technology and Business University, China*
[c] *School of Computer and Electronic Information, Nanjing Normal University, China*
[d] *Department of Computer Science, Virginia Tech, USA*

ARTICLE INFO

Article history:
Received 16 February 2022
Received in revised form 11 March 2022
Accepted 13 March 2022
Available online 23 March 2022

Keywords:
Metaverse
Virtual reality/Augmented reality
Computer graphics
User interaction
Immersive visualization

ABSTRACT

The metaverse is a visual world that blends the physical world and digital world. At present, the development of the metaverse is still in the early stage, and there lacks a framework for the visual construction and exploration of the metaverse. In this paper, we propose a framework that summarizes how graphics, interaction, and visualization techniques support the visual construction of the metaverse and user-centric exploration. We introduce three kinds of visual elements that compose the metaverse and the two graphical construction methods in a pipeline. We propose a taxonomy of interaction technologies based on interaction tasks, user actions, feedback and various sensory channels, and a taxonomy of visualization techniques that assist user awareness. Current potential applications and future opportunities are discussed in the context of visual construction and exploration of the metaverse. We hope this paper can provide a stepping stone for further research in the area of graphics, interaction and visualization in the metaverse.

1. Introduction

The word "metaverse" was first coined in the 1992 science fiction novel named *Snow Crash*, written by *Neal Stephenson* (Joshua, 2017). The novel depicts people in a virtual reality world competing against each other for social status by controlling their digital avatars. Over the next 30 years or so, the concept of metaverse has appeared in books and movie shows. At this stage, the concept of the metaverse is ambiguous and is more understood as a virtual world parallel to the real world. For example, the film Ready Player One (Spielberg et al., 2018) described a virtual world in which everyone could customize their avatars and explore freely through multimedia techniques (VR/AR, etc.). It was considered a classic film that interpreted the concept of the metaverse. However, it is not enough to rely on multimedia technologies. Metaverse needs to be able to provide users with a more realistic experience and rich activities, which requires more advanced technologies to support metaverse's construction and user-centric exploration. Recently, some scholars defined metaverse from an aspect of a comprehensive technical architecture. Lee et al. defined metaverse as a 3D virtual cyberspace blending the physical and digital world, facilitated by the convergence between the Internet and Web technologies and Extended Reality (XR) (Lee et al., 2021a). Duan et al. also categorized the related technologies of metaverse into three levels: infrastructure, interaction and ecosystem (Duan et al., 2021). Different from their overall perspective, we do not discuss the infrastructure techniques like Blockchain, Network and Edge computing. We focus on the critical technologies that support visual construction and user exploration, including Graphics, Interaction and Visualization.

The visual construction of metaverse is based on the graphical techniques that build the integrated world combining the physical and virtual world, including the 3D construction of scenes, non-player characters (NPCs) and player characters (Avatar). Interactive technology enables users to operate visual elements, explore freely in the metaverse and provide an immersive experience. To improve users' awareness of the virtual world, supplementary instructions and guidance are required. Visualization can provide such guidance by processing the data in the metaverse and presenting it to users in an appropriate form. The development of these technologies makes metaverse more realistic and interesting for users to perceive and explore.

In this work, we propose a framework for visual construction and exploration of metaverse (Fig. 1), where we start with what can be seen in the metaverse and what people can do in it, summarizing (1) the visual elements (scenes, NPCs and Avatar) in metaverse environment and how the graphical techniques used

---

* Corresponding author.
E-mail addresses: yuhengzhao@fudan.edu.cn (Y. Zhao), simingchen@fudan.edu.cn (S. Chen).

https://doi.org/10.1016/j.visinf.2022.03.002

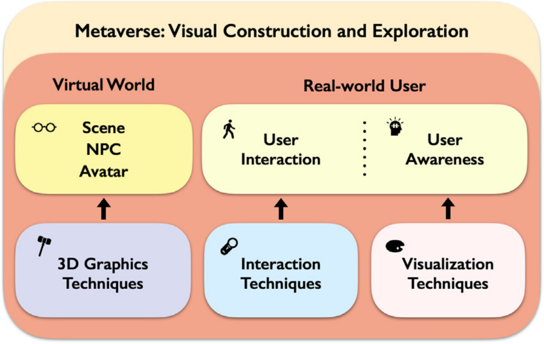


Fig. 1. The framework of this survey. The essential visual elements users can see compose the metaverse environment. The explorative behaviors that users can conduct in such an environment. Three key technologies fuel the visual construction and exploration of the metaverse.

to create them, and (2) exploration of metaverse (interaction and awareness) and how interaction and visualization techniques can support these behaviors. We discuss the current applications and future opportunities in the context of our framework. To our best knowledge, this survey serves as the first effort to offer a view of visual construction and exploration of the metaverse. The detailed contributions are as followed.

- A **framework** of visual construction and user exploration of metaverse from the perspectives of graphics, interaction and visualization.
- A visual construction **pipeline** that consists of two methods and three phases supported by graphics techniques and two **taxonomies** of interaction and visualization techniques for user interaction and awareness when exploring in the metaverse.
- A set of **research challenges and opportunities** derived from our review of techniques and applications.

We hope this paper can provide a stepping stone for further research in the area of visual construction and user exploration of the metaverse.

2. Overview

As the metaverse is still in its infancy of development, no work has been done to systematically summarize the technical framework for its complete visual construction and exploration, nor have graphics, interaction and visualization been explored separately from the context of the metaverse. The work most closely related to ours investigates the current state of research in their respective fields.

In terms of computer graphics, many scholars have introduced how to build 3D models. Some of them proposed software-based authoring processes (Tang and Ho, 2020) and automated generation to reduce processes (Freiknecht and Effelsberg, 2017). Some introduced 3D reconstruction of real objects (Intwala and Magikar, 2016). However, there lacks a comprehensive introduction to visual elements construction in the metaverse. We summarize three kinds of elements and propose a pipeline that consists of the above two methods and compare the construction differences between different elements. As the elements in metaverse are often dynamic and interactive, we investigated 3D animation as a part of the pipeline.

Surveys on interactions specifically related to metaverse are limited as well. Some of them summarize from different sensory channels, such as haptic solutions (Bouzbib et al., 2021). Besides, there are surveys summarizing interaction methods for certain devices, such as smart glasses (Lee and Hui, 2018), optical head-mounted displays (Grubert et al., 2017), virtual reality headset (Kelly et al., 2021). Rather than focusing on interaction methods via certain sensory channels or concentrating on interaction methods supported by particular devices, we try to summarize interactions methods used in metaverse from perspectives of interaction tasks, user action and feedback via various channels and devices.

From the perspective of visualization, many scholars have summarized immersive visual analytics from the data types and environments (AR/VR) (Kraus et al., 2021), interactions for visualization (Fonnet and Prie, 2019; Besançon et al., 2021) or its future challenges (Ens et al., 2021a). However, there lack discussions about the connections and challenges in the metaverse. A basic connection is to help users perceive the environment, for example, the map visualization of spatial data is used for navigation. However, with the development of the metaverse, there will be more complex visualizations that need to be investigated to help perceive the environment or further application to understanding and analysis. Therefore, we will discuss visualization from data types, visualization view, position, interactions, and environments, by fully considering the diversity, shareability, and user-centered experience of the metaverse.

This paper serves as the first effort to summarize the technical framework for visual construction and exploration in the metaverse. In the next section, we will first introduce the visual elements that compose the metaverse and how they are constructed in terms of design-based and physical-based construction methods with a pipeline. In Section 4, we will introduce the two taxonomies of interaction that support user interaction, and visualization techniques that support user awareness in exploring the metaverse. In Section 5, we will summarize current metaverse applications (including virtual social, smart city, medical and games) in terms of visual construction and exploration. We will discuss the current limitations and future research opportunities derived from our analysis in Section 6.

3. Virtual world construction

The virtual environment of metaverse blends physical and digital, which consists of various scenes, non-player characters (NPCs) and player characters (Avatars). **Scene** refers to diversified virtual spaces, such as virtual campus (Duan et al., 2021) or virtual museum (Beer, 2015). **NPC** is an object that cannot be controlled by the player but has an important role in the game itself so that it makes the world in the game feel alive (Warpefelt and Verhagen, 2015). **Avatar** refers to the digital representation of players in the metaverse, where players interact with the other players or the computer agents through the avatar (Davis et al., 2009). The creation of these objects is based on computer graphics techniques. The scene, NPC, and PC differ in the detail of their creation because they focus on different features. In this section, we will introduce a pipeline and two ways for constructing virtual models in metaverse: physical-based or design-based construction, and compare the construction of scenes, NPCs and Avatars.

3.1. Visual construction pipeline

The pipeline can be divided into three stages: initialization, modeling and rendering, and animation (Fig. 2).

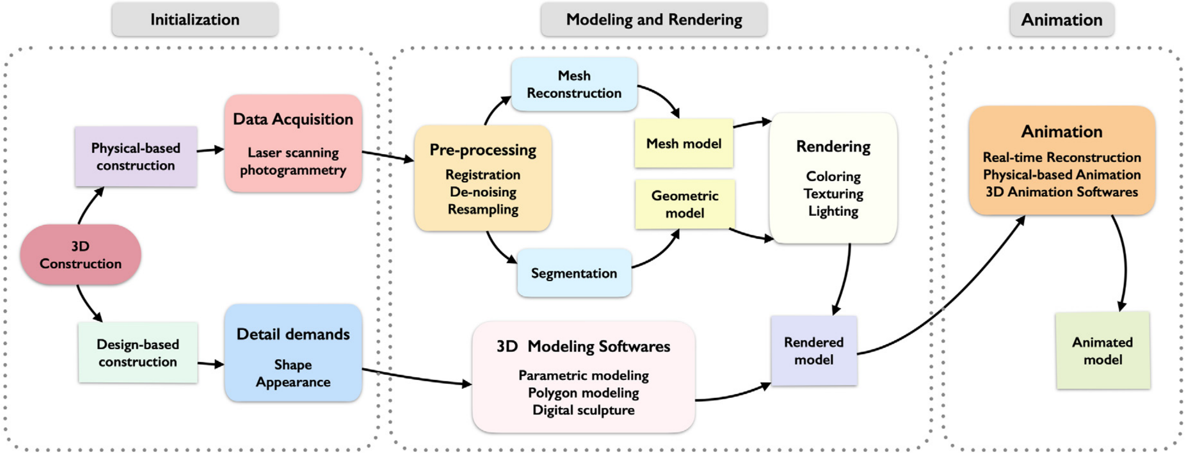


Fig. 2. The two ways to construct a virtual scene in metaverse: physical-based construction and design-based construction, and the construction process can be divided into three steps: Initialization, Modeling and Rendering, and Animation.

3.1.1. Physical-based construction

One way to build 3D models is the physical-based construction by using 3D measurement methods, primarily laser scanning and photogrammetry, to create digital twin models, termed 3D reconstruction (Deng et al., 2021). Recently, the global trend toward Virtual Reality (VR) and Augmented Reality (AR) has increased the demand for creating high-quality and detailed photorealistic 3D content based on real objects and environments. One major difficulty is how to restore high fidelity with collected data.

Data acquisition is the first stage of 3D model reconstruction in the metaverse, which determines how realistic the model is. With the rapid development of 3D acquisition technologies, 3D sensors are becoming increasingly available and affordable. At present, there are two popular and convenient point cloud acquisition methods: laser scanners (LiDARs), RGB-D cameras (such as Kinect, RealSense and Apple depth cameras). 3D data acquired by these sensors can provide rich geometric, shape and scale information (Guo et al., 2021).

Preprocessing refers to the registration, denoising and resampling of the obtained point cloud model (Ma and Liu, 2018). Notably, which of three steps are required depends on the application requirements and the quality of data.

Normally, an object usually needs many site scans, which requires multi-site data registration, that is, the data of each site is converted to the same coordinate by using a cloud-to-cloud alignment tool (Intwala and Magikar, 2016). The algorithm typically used for registration is the ICP algorithm (Khaloo and Lattanzi, 2017), which estimates the rigid transformation between two point clouds iteratively by minimizing the distance. Recently, the developments of optimization-based methods and deep learning methods have improved registration robustness and efficiency (Huang et al., 2021; Perez-Gonzalez et al., 2019; Wang and Solomon, 2019). Point clouds after registration incorporate the points from the objects of interest and the noise points. Therefore, de-noising is normally a needed step. An effective method is random sample consensus (Liu et al., 2016). Finally, the amount of data after registration will be extremely large and makes overlapped regions denser, which will reduce the efficiency of subsequent processing. Resampling is necessary to solve such problems. In practice, three commonly used approaches are random, uniform, and point-spacing (Son et al., 2015).

Modeling is the core process in visual construction. The prepossessed point cloud to build contains rich information of the target object, but it still needs to be converted into a 3D model represented by basic geometric shapes such as planes, surfaces and cuboids, etc. There are major two modeling methods: geometric modeling and mesh reconstruction.

One modeling method is to generate geometric models by point cloud segmentation and modeling. A segment is characterized by a planar cluster because it is a set of points located within a given threshold distance about a calculated plane. After segmentation, the segmented planar clusters are used to extract the main contours of architectural components. The region growth (Pu et al., 2006) and RANSAC (Liu et al., 2016) are two efficient segmentation methods (Nguyen and Le, 2013). Another commonly used method is to generate a mesh model since it can be used for complex surface modeling, which can be generally established by constructing a triangular net or NURBS (Non-uniform rational basis spline). Nevertheless, the two approaches can also be combined to create semantically segmented 3D mesh models (Leotta et al., 2019).

Rendering is an important step for endowing virtual objects' real features and presenting them in front of our eyes. Rendering usually includes coloring, texturing and lighting. Since laser scanning lacks color information that is required in many applications. Therefore, a hybrid 3D reconstruction based on images and scan data is adopted to colorize the point cloud data. For example, Ma et al. proposed a differential framework for freestyle material capturing (Ma et al., 2021). This technique can greatly improve the sampling efficiency and restore the material properties with higher precision, which supports the creation of more realistic "virtual things" in the metaverse.

Physical-based animation has emerged as a core area of computer graphics and is widely applied in many areas such as film, game and virtual reality (Bargteil et al., 2020). In general, physical-based animation is a method to simulate and animate the dynamic changes of objects based on physical rules or motion control theory. For example, the animation of rain falling on leaves, the swing of leaves under the wind and the floating animation of leaves on the water can be realized through physical simulation. These physical animations of objects in a virtual scene will make people wander the metaverse as if they were in the real world. When applied in the virtual reality environment, the research focuses more on how to generate animations based on interactive models (Llobera et al., 2021).

3.1.2. Design-based construction

Another way to create a 3D model is using specialized 3D modeling software.

Detail demands in the start of design-based construction from the designer's ideas (e,g. conception of shape and appearance) and iterates during the drawing process. The advantage of this approach is that we can design entirely based on imagination, for example with objects designed with futuristic technology and unheard of fantasy creatures, or things that cannot be seen or scanned up close.

Modeling and rendering can be implemented by a number of modeling software available and they can be divided into three categories. First, parametric 3D modeling or CAD (computer-aided design) such as AutoCAD[1] and SketchUp,[2] is the preferred method for engineers and designers to create models by setting the parameters as the real thing: materials, weight, etc. Second, unlike CAD modeling, polygon models are more concept-oriented than measurement-oriented. The popular software is 3Ds Max,[3] Maya,[4] Blender,[5] etc. Third, digital sculpture modeling software such as ZBrush,[6] requires more artistic skill than polygon modeling. The modeling software can be combined with engines (Unity3D[7] or Engine[8]) to construct the virtual environment.

Design-based animation is achieved by computer graphics software to make objects appear to move in 3D space. The designers use such software to construct the simple object first, followed by rigging. The animator places rigs at strategic points to make it appear to be moving. Some of the common modeling software we mentioned above can support animation creations. For example, Blender is an all-rounder, capable of handling a full line of tasks from 3D modeling and animation to video editing. Recently, some interactive controller-based animating are proposed. For example, AnimationVR is proposed as a plugin for Unity, which supports beginners to easily create animations in virtual reality (Vogel et al., 2018).

3.2. Scene

In the metaverse, users will see many colorful and realistic scenes. The construction of the scene focuses more on the realism of architecture and layout.

In the modeling process, mesh models are more suitable for reconstructing some buildings with complex surface shapes, such as murals and statues. For example, Leotta et al. reconstructed urban buildings combining segmentation methods into mesh models (Leotta et al., 2019, Fig. 3a). Navarro et al. used the mesh method to reconstruct indoor rooms into virtual worlds (Navarro et al., 2017, Fig. 3b). Some researchers also studied how to make remote multiple clients realize real-time reconstruction access in the virtual environment. SLAMCast achieved a practical client–server system for real-time capture and many-user exploration of static 3D scenes (Stotko et al., 2019, Fig. 3c).

3.3. Non-player character (NPC)

In metaverse, people can interact with NPCs by communication and gestures, etc, and can even befriend AI-driven NPCs (Duan et al., 2021). These vivid characters are created by computer graphics technologies. The creation process is similar to the scene construction, but the modeling and animations of NPCs mainly focuses on how to make them more like real people or animals, both in appearance, in behavior and intelligence.

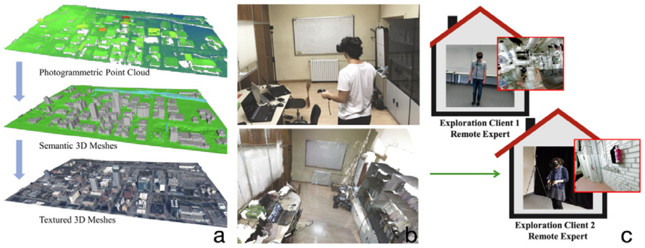


Fig. 3. Examples of scene construction. (a) Urban 3D reconstruction from multi-view satellite imagery (Leotta et al., 2019). (b) A room scene 3D reconstruction with an RGB-D camera (Navarro et al., 2017). (c) A remote client–server system for real-time reconstruction (Stotko et al., 2019).

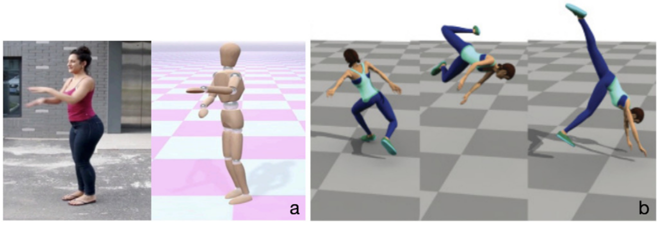


Fig. 4. Examples of construction and animation of NPCs. (a) A deep neural network that directly reconstructs the motion of a 3D human skeleton from monocular video (Shi et al., 2020). (b) Adapted Q-networks to create a control system that can respond to disturbances and allows steering and other user interactions (Liu and Hodgins, 2017).

Modeling and rendering for characters in metaverse pay more attention to the features of people, such as skin, hair and clothing, etc. This has high requirements for modeling and rendering technology. Many researchers have proposed relevant methods, for example, Lyu et al. (2020) proposed the first CNN-integrated framework for simulating various hairstyles.

The research goal of generating animation for NPCs is to model the action, behavior and decision-making process of humans and animals. Li et al. summarized seven commonly used animation methods (Li et al., 2021), of which finite state machines (FSM) are the most straightforward and widely adopted model for NPC to respond to players' behavior (Lau and Kuffner, 2005). For example, Dehesa et al. proposed a framework based on a state machine to generate responsive interactive sword fighting animations against player attacks in virtual reality (Dehesa et al., 2020). Some scholars also used deep learning techniques to reconstruct motions from videos (Shi et al., 2020, Fig. 4a) and reinforcement learning has combined to enable NPCs to learn automatically from the interactive experience of their surroundings (Liu and Hodgins, 2017, Fig. 4b).

3.4. Player character (avatar)

Compared to the other two elements, the construction of the avatar is more user-defined, with many features coming from users, in both the 3D modeling and animation creation.

Traditionally, users can modify and edit the appearance of their avatars with many options. To mimic the users' real-life appearances, there occurs some applications that allow users to scan their physical appearance, and subsequently choose their virtual outfits. Although design-based avatars have many improvements in their sense of realism, which are still carton-like. During various social activities in the metaverse, the details of the avatar's face (Wei et al., 2004) and the micro-expression (Murphy, 2017), the whole body (Kocur et al., 2020) could impact the user perceptions. Therefore, to improve the senses of realism, many

---

[1] https://web.autocad.com/.

[2] https://www.sketchup.com.

[3] https://www.autodesk.com.hk/products/3ds-max/.

[4] https://www.autodesk.com.hk/products/maya/.

[5] https://www.blender.org/.

[6] https://pixologic.com/.

[7] https://unity.com/ or Unreal.

[8] https://www.unrealengine.com/.

reconstruction technologies are developed for highly realistic 3D faces and bodies.

Reconstruction of the face is an important part, which is usually based on 3D Morphable Model (Booth et al., 2016) generated by Principal Component Analysis (PCA) algorithm, but the performance is limited by data quality, difficult to express facial details. In recent years, with the development of deep learning technology, a series of Generative Adversarial Networks (GAN) and SDF have emerged and achieved a higher degree of realism. For example, Luo et al. (2021) used StyleGAN to generate highly photorealistic 3d face template.

The research of 3D body has been widely done, e.g. the 3D body micro-neural rendering based on different data types (Wu et al., 2020), the high realistic human dynamic geometry and texture reconstruction based on simple input (Liu et al., 2021a). When scanning or capturing the full-body, hands usually occluded due to the small size. To cope with such low-resolution, occlusion problems, researchers have proposed many models to express human hands. The most famous model is MANO (Romero et al., 2017), which can well adapt to the deep learning network.

The animation of avatars is commonly generated by user manipulation, e.g. interactions with controllers or real-time tracking (Genay et al., 2021). On one hand, through interaction, users can share actions with their avatars. On the other hand, since control through inputs that are not physically representative of the user's face, hands or body, might therefore not play in favor of strong ownership illusions, real-time tracking technologies that can provide a mapping of user movements have been the important tool for animation generation. For example, Saragih et al. proposed a real-time system that achieved convincing photorealistic avatar and faces animation from a single image (Saragih et al., 2011). Mueller et al. combined a convolutional neural network with a kinematic 3D hand model, which addressed the highly challenging problem of real-time 3D hand tracking (Mueller et al., 2018).

4. Real-world user in metaverse

The metaverse is a user-centric application by design. As such, every component of the multiverse should place the human user at its core (Lee et al., 2021a). Therefore, the key to exploring the metaverse is to provide a good experience for users, including reasonable and effective interaction, and accessible visual guidance or hints to aid in rapid awareness, comprehension and analysis. This section aims to describe state-of-the-art interactive and visualization techniques that can support user interaction and awareness in the metaverse.

4.1. User interaction

To a given interaction task, an effective and complete interaction process can be realized by the user action (input) and feedback (output) from various devices. To realize the whole process, various sensory channels are combined (Fig. 5).

Interaction tasks refer to various ways that enable users to contact, control or influence the metaverse. There are various kinds of classification methods, for example, Raaen and Sørum divided interactions tasks into menus, locomotion and interaction (Raaen and Sørum, 2019). In the context of the metaverse, we decompose interaction tasks into three elementary manipulation processes: navigation, contact and editing.

User Actions refer to behaviors that users can achieve through various sensory channels, such as gaze and gestures through body language channel, which is the input of an interaction task.

Feedback refers to the responses from devices to a user action, such as changes of view in smart glasses and forces generated by controllers, which requires various sensory channels of users to participate at the same time. Feedback can be decomposed into 5 types: visual channel, acoustic channel (auditory channel), haptic channel, olfactory channel and gustatory channel based on different sensory channels used.

Sensory channels refer to the perceptual senses used by devices during an interaction task. We divide channels in two ways by different tasks: converting user action into digital input and converting digital output as feedback to users.

4.1.1. Interaction tasks

Navigation refers to user operations that result in view changes in the metaverse. Navigation tasks can be decomposed into navigation by geographic cues and navigation by non-geographic cues. Mainly there are four different kinds of navigation by geographic cues: real walking with the displacement of users, panning using controllers, changing viewpoint with head movement and pointing and teleport using controllers. Users can also be navigated by non-geographic cues such as navigation by query, by tasks, or other specified movements (Jankowski and Hachet, 2013).

Contact refers to the ability to touch and feel objects by controllers in the environment. Contact tasks can be divided into two different ways: direct contact and indirect contact. Direct contact refers to interactions by using controllers as bare hands, fingertips or any body parts, and are highly similar to movements in the real world when direct touching happens. Indirect touching refers to interactions using certain things in the metaverse. Under that circumstance, devices, such as hand controllers, are considered as continuity to the user's hands. For instance, the user may need to use the controller as a knife to touch and feel the organs of the body when operating surgeries in the metaverse.

Editing refers to the interaction process that involves changing either properties, position or orientation of any objects in metaverse. The editing process includes two phases: selection and manipulation.

4.1.2. Various channels of user action

Body language and sonification channel can be decoded from gaze, which requires eye-tracking technology, or from gesture, which can be recognized with gesture recognition-based sensors, or from sonification, which can communicate with objects in metaverse.

Haptic channel consists of two types: tactile and kinesthetic, which combine the sense of touch used in the interaction process. The tactile cues are developed through the skin, while the kinesthetic ones come from proprioception and are through the muscles and the tendons (Bouzbib et al., 2021).

Brain signal channel also can be used as an input, which is widely utilized in brain–computer interface systems. Electro-Encephalo-Graphy (EEG), electromyography (EMG) biopotentials, SSVER, P3 EP and many other signals can be used as control inputs to several specific tasks in the metaverse (Friedman et al., 2007). For example, the brain can let users change the camera position in metaverse toward the left or right by using two different brain signals, such as left- or right-hand motor imagery (MI) or two steady-state visual-evoked potentials (SSVEPs) at different frequencies (Lécuyer et al., 2008).

4.1.3. Various channels of feedback

Visual channel and auditory channel refers to feedback generated based on the vision and the sense of acoustic. The visual feedback is related to changes generated by user action, displayed in a stereoscopic viewpoint via smart glasses or headsets. The auditory feedback can be further decomposed into two different

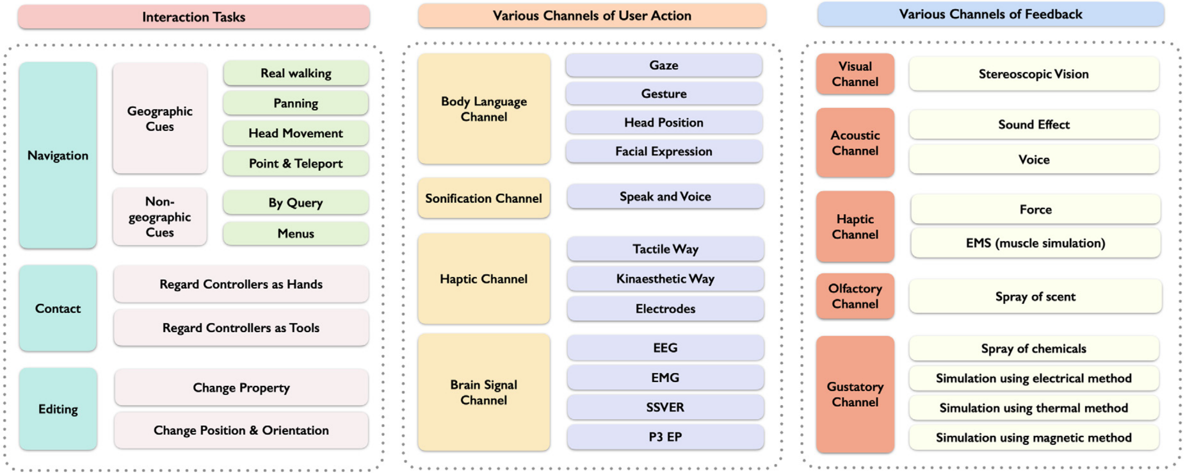


Fig. 5. The taxonomy of interaction tasks, user action (input) and device feedback (output) with various sensory channels that we summarized in this paper.

types: sound effect, which can make users better immerse themselves into metaverse, and voice which is key acoustic feedback to allow users to communicate with an avatar in metaverse with more sense of reality.

Haptic channel refers to feedback generated based on the sense of touch. Users can feel different materials, textures, temperatures or feel shapes and patterns through their fingertips, and can perceive stickiness, smoothness, pressure, vibration or friction. Feedback mentioned above involves multiple force types: tension, traction, reaction, resistance and impact, which help enhance the user experience in the metaverse. Gloves or exoskeletons constrain the user's hands for simulating shapes or stimulate other haptic features such as stiffness, friction or slippage (Bouzbib et al., 2021).

Olfactory channel refer to feedback generated based on the sense of smell. The major way of olfactory feedback is dispersing specific scents to match the requirements of certain interaction tasks. Olfactory displays are widely used to disperse scent, which synchronizes odorants from a digital description. The feedback from the olfactory channel can better immerse users in the metaverse. Generally, an Olfactory display consists of a palette of odorants, a flow delivery system and a control algorithm that determines the mixing ratios, concentration and timing of the stimulus (Richard et al., 2006).

Gustatory channel refers to feedback generated based on the sense of taste. There are mainly two ways of actuating the sense of taste, using the chemical combination to spray into the user's mouth or to spray to an area where the user can lick with tongue, or using digital taste actuation technology to produce various taste sensations. Devices can be developed to fulfill the process of gustatory stimulation with or without chemicals, for example, the "Virtual cocoon" can spray flavors directly into the user's mouth, and the "Food Simulator" uses chemical and mechanical linkages to simulate food chewing sensations by providing flavoring chemicals, biting force, chewing sound, and vibration to the user (Ranasinghe et al., 2011). Applying thermal stimulation on the tongue and stimulating the TRPM5 (Transient receptor potential cation channel subfamily M member 5) taste channel can enhance the flavor of sweet, bitter, and umami tastes (Karunanayaka et al., 2018).

4.1.4. Collaborative interaction

Generally, collaborative interaction can be decomposed into three categories: communication, joint navigation and collaborative editing.

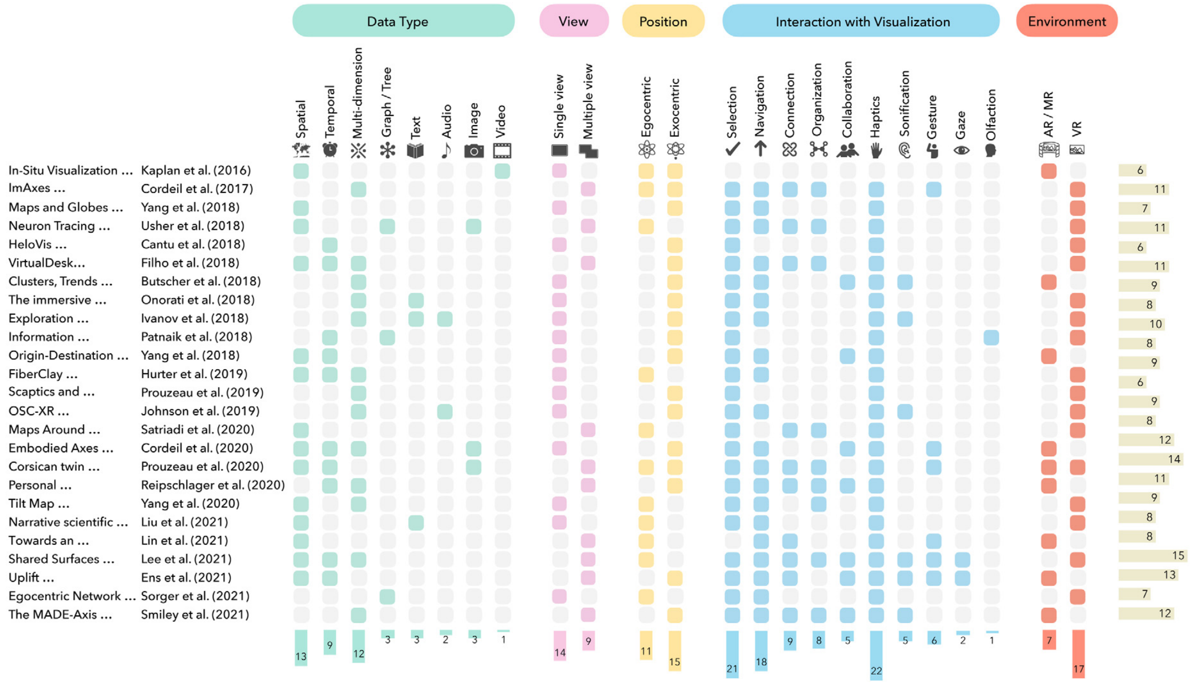
Communication can be further categorized as gestural communication, verbal communication and other ways of using body language. Users can transfer information by gestures, such as tracing along the boundaries of objects or simply pointing at each other (Beck et al., 2013). Certain gestures can be utilized as established interaction methods according to design mechanism, taking advantage of users' social intuition and communicative skills (Roth et al., 2015). In verbal communication, the combination ratio of intonation, speech speed and sound volume of an avatar, if not transmitted directed from a user's voice, are frequently adjusted to affect the level of effectiveness during the social interaction process (Eynard et al., 2015). Other body languages, such as body motion, facial expressions can be tracked, transmitted and represented via virtual avatars in the metaverse (Roth et al., 2017). Devices and algorithms may be required for high rapport and better understanding when intercultural communications happen. For instance, in a conversation between a Japanese and a German, the bow may need to translate into the handshake displayed by the corresponding avatar (Roth et al., 2015).

Joint navigation refers to multi-user operations that result in view changes in the metaverse. In a complete joint navigation process, four different techniques are required: forming navigational groups by multi-users, distributing navigational responsibilities, performing navigation tasks together and ending joint navigation by splitting up (Weissker et al., 2020).

Collaborative editing refers to the interaction process that involves changing either properties, position or orientation of any objects by multi-users in the metaverse. Users can manipulate objects together, and any user-crested element can be kept, changed or moved by other users (Greenwald et al., 2017).

4.2. User awareness

In the metaverse, users are exposed to the environment composed of different data. The perception of the complex environment through the processing of data is needed, and in further scenarios, especially with understanding and analytical needs (VR meeting, etc.), a detailed analysis of the data is required. These perception, comprehension and analysis scenarios are considered as user awareness when exploring in the metaverse. In this section, we summarize existing advanced research from five dimensions (Fig. 6) to help guide the related research in the metaverse. To build our taxonomy, we collected papers published on leading conferences and journals in Visualization and HCI, including Vis, TVCG, Eurovis, CHI, etc. We chose the papers if



		Data Type								View		Position		Interaction with Visualization										Environment		
		Spatial	Temporal	Multi-dimension	Graph / Tree	Text	Audio	Image	Video	Single view	Multiple view	Egocentric	Exocentric	Selection	Navigation	Connection	Organization	Collaboration	Haptics	Sonification	Gesture	Gaze	Olfaction	AR / MR	VR	
In-Situ Visualization …	Kaplan et al. (2016)	●							●	●		●	●											●		6
ImAxes …	Cordeil et al. (2017)			●							●	●	●	●	●	●	●		●		●				●	11
Maps and Globes …	Yang et al. (2018)	●								●			●	●	●				●						●	7
Neuron Tracing …	Usher et al. (2018)	●			●			●			●	●		●	●	●	●		●						●	11
HeloVis …	Cantu et al. (2018)		●							●			●	●					●						●	6
VirtualDesk…	Filho et al. (2018)	●	●	●							●		●	●	●	●	●		●						●	11
Clusters, Trends …	Butscher et al. (2018)			●						●			●	●	●			●	●	●				●		9
The immersive …	Onorati et al. (2018)			●		●				●			●	●	●				●						●	8
Exploration …	Ivanov et al. (2018)			●		●	●			●			●	●	●				●	●					●	10
Information …	Patnaik et al. (2018)		●		●					●			●	●					●				●		●	8
Origin-Destination …	Yang et al. (2018)	●	●							●			●	●	●			●	●					●		9
FiberClay …	Hurter et al. (2019)	●	●	●						●		●		●	●				●						●	
Scaptics and …	Prouzeau et al. (2019)			●						●			●	●					●						●	6
OSC-XR …	Johnson et al. (2019)			●			●			●			●	●	●				●	●					●	9
Maps Around …	Satriadi et al. (2020)	●									●	●		●		●	●		●						●	8
Embodied Axes …	Cordeil et al. (2020)	●	●	●				●		●			●	●	●			●	●		●			●		12
Corsican twin …	Prouzeau et al. (2020)	●	●					●			●	●	●	●	●	●	●		●		●			●	●	14
Personal …	Reipschlager et al. (2020)		●	●							●		●	●	●	●	●	●	●					●		11
Tilt Map …	Yang et al. (2020)	●		●						●		●		●	●		●		●						●	9
Narrative scientific …	Liu et al. (2021)	●				●				●		●		●	●				●						●	8
Towards an …	Lin et al. (2021)	●									●	●			●	●			●		●			●		8
Shared Surfaces …	Lee et al. (2021)	●	●	●							●	●		●	●	●	●	●	●	●	●	●			●	15
Uplift …	Ens et al. (2021)	●	●								●		●	●	●	●		●	●	●	●	●		●		13
Egocentric Network …	Sorger et al. (2021)				●					●		●		●	●				●						●	7
The MADE-Axis …	Smiley et al. (2021)			●							●		●	●	●	●	●	●	●	●				●		12
		13	9	12	3	3	2	3	1	14	9	11	15	21	18	9	8	5	22	5	6	2	1	7	17	

Fig. 6. We reviewed 20 key articles from 2016 to 2021 and summarized the taxonomy of immersive visualization from 5 dimensions. The boxes with different colors represent Data Visualization (🟩), View (🟪), Position (🟨), Interaction for visualization (🟦), and Environment (🟥). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

they discuss the metaverse from perspectives of visualization, immersive environment and immersive interaction. Eventually, we found 25 key publications and organized them into a meaningful taxonomy for our main analysis.

4.2.1. Data type

Different data types often correspond to different scenarios and exploration tasks in the metaverse, so there are different data visualizations and different ways of interaction, etc. We find that the currently available types of immersive visualization can be divided into spatial and tabular data, multidimensional and rational data, medium data.

Spatial and tabular data generally refers to common data such as spatial data (e.g. geographic, scientific) (Lin et al., 2021; Hurter et al., 2018), temporal data (Cantu et al., 2018; Prouzeau et al., 2020). For example, some scholars proposed immersive map visualizations (Yang et al., 2018b, 2020; Satriadi et al., 2020; Yang et al., 2018a) or visualized geographic data (White and Feiner, 2009), which are essential for navigation in metaverse.

Rational data includes multidimensional data (Filho et al., 2018), graph/tree structure data, etc, and often corresponds to higher analytical needs in the metaverse, such as virtual analysis workshops or presentations. For example, ImAxes is an interactive multi-dimensional visualization tool for understanding such rational data, which is a basic technology for future analysis in the metaverse (Cordeil et al., 2017).

Multimedia data (text, audio, image and video) that often appears in metaverse need to be perceived in an appropriate way. Due to the complexity and unstructured feature, the related work is still relatively scarce. Currently, text visualization is primarily used to help users explore by generating descriptions to aid comprehend (Ivanov et al., 2018; Liu et al., 2021c) or help analysts to understand semantic features (Onorati et al., 2018). Audio data also can be visualized in an immersive environment for music interaction (Johnson et al., 2019), storytelling (Latif et al., 2021; Kaplan et al., 2016) visualizing kinetic metrics of foot pedals in video data to aid motor training.

4.2.2. View

A visualization may consist of a single view or multiple views. A **single view** refers to the representation of a set of data in a single window. **Multiple views** refer to any instance where data is represented in multiple windows. While simple exploration can be implemented with a single view, many complex tasks such as multiple or data types, visual comparison tasks benefit from multiple views. For example, users can perform map exploration search, comparison and route-planning tasks by a multi-view map visualization (Satriadi et al., 2020). Therefore, connection and organization interactions between multiple views are required for such tasks. Such connection includes simultaneous updating or highlighting corresponding information between views, etc. The organization is to arrange the views, e.g., in a gallery or sequence (Batch et al., 2019).

4.2.3. Position

To provide a good perception for the user, the position of the view(s) relative to the user also needs to be adjusted with an egocentric or an exocentric perspective (Ens et al., 2014). An **egocentric** position means that view is arranged regarding the user's position and around the user. An **exocentric** position means that the view(s) is arranged regardless of the user's position. By reviewing the literature, we found that position type is related to task accuracy and task mode.

The exocentric position is more accurate for search and judgment tasks since being outside of the data in exocentric affords a full overview, which is less fatiguing and is easier to use in existing analyst work-spaces (Wagner et al., 2021). By contrast, being inside the data in egocentric allows the observation of details through spontaneous exploration. Yang et al. also found that exocentric globes are more accurate and faster with overview observations (direction or area), while egocentric is more suitable for observing small variations in detail (Yang et al., 2018b). Therefore, exocentric and egocentric are suitable for different tasks. When exploring details of a large network, a simple egocentric

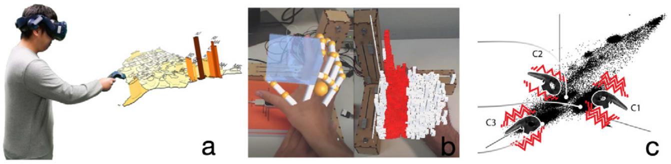


Fig. 7. Examples of interaction for visualization. (a) Tilting a 3D prism map to 2D chart by controllers (Yang et al., 2020). (b) Comparing tangible axes and gesture interaction (Cordeil et al., 2020). (c) Vibration feedback encodes the density in a 3D scatter visualization (Prouzeau et al., 2019).

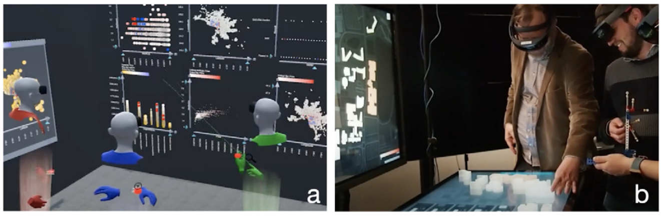


Fig. 8. Examples of collaborative interaction. (a) The FIESTA system for collaborative immersive analytics in VR (Lee et al., 2021b). (b) The UpLift system uses tangibles widgets and the Hololens to visualize and understand building energy (Ens et al., 2021b).

interface can considerably improve the efficiency (Sorger et al., 2021).

There are two kinds of task modes: exploration and presentation. Batch et al. conducted an evaluation (Batch et al., 2019) that used the ImAxes immersive visualization system (Cordeil et al., 2017). The results indicated that participants placed visualizations egocentrically and in close-range exploration tasks since it is more efficient to use local space around the user. Conversely, participants used more space to arrange visualizations in an exocentric way when presenting insights to others in a collaborative setting. Based on such evaluation, Satriadi et al. also adopted an egocentric position that allows users to create large hierarchies of multiple maps at different scales and arrange them in 3D space (Satriadi et al., 2020).

4.2.4. *Interaction for visualization*

Different from Section 4.1, this section focuses on the interactions in the exploration of different data visualizations in the metaverse. For example, we add connection and organization for multi-views, which refer to the capability to support the coordination and arrangement of multiple views. Some channels, such as brain signal and gustation, have limited integration with data visualization and are therefore not explicitly listed.

Interaction for visualization with various channels is implemented by using haptic controllers or other senses to perform various manipulations.

In terms of haptic controllers, Yang et al. introduced tilting for transitioning between a 2D and 3D map (Yang et al., 2020, Fig. 7a). When using controllers for navigation, zoom or overview is better than standard locomotion alone (Yang et al., 2021). Moreover, some scholars explored tangible widgets in immersive visualization (Smiley et al., 2021). For example, embodied axes (Cordeil et al., 2020) utilized three orthogonal arms to represent data axes for selection and navigation (Fig. 7b). Butscher et al. explored parallel coordinates for multi-dimensional data using a touchable desktop (Butscher et al., 2018). These efforts tend to have better accuracy since they are fixed in physical space.

Other sensing channels can also help interactions, which can be freer and more in tune with human behaviors, but with a corresponding accuracy reduction. Body languages like gestures are widely used, for example embodied axes support mid-air gestures for selection (Cordeil et al., 2020, Fig. 7). The sonification channel is usually used as a voice input command, the combination of natural language techniques allows the machine to help users interact with the visualization (Liu et al., 2021b). However, such methods with immersive visualization in metaverse still need to be explored.

Interaction for visualization with feedback refers to using various channels to perceive the results of an interaction. Haptic feedback often takes the form of vibration, which improves the accuracy of the user's perception. Prouzeau et al. encoded data density as vibration intensity and provided feedback through the haptic controller (Prouzeau et al., 2019, Fig. 7c). The vibration can improve user performance for identifying void regions, which is helpful for users to perceive with the assistance of data visualization, especially unknown areas in the metaverse. Usher et al. applied vibration with selection, which enables users to have a stronger perception of accuracy (Usher et al., 2018).

Sonic feedback tends to be sound effects that fit the context of the exploration. Ivanov et al. implemented a novel unit visualization with many 3D avatars, and they proposed the opportunity to leverage audio mappings to create a strong association between visual and audio elements, such as unique sounds, pitches, or speech (Ivanov et al., 2018). Olfaction can also used to transmit information (Batch et al., 2020). For example, in a 2D network graph, Patnaik et al. chose smell-color combinations such as lemon-orange, leather-red, coconut-white to represent nodes (Patnaik et al., 2018).

Collaboration interaction for visualization allows multiple users to simultaneously explore a meta-cosmic environment with the help of data visualization.

From a perspective of environment, most of them are implemented in an AR environment (Ens et al., 2021b; Lee et al., 2021b; Butscher et al., 2018; Reipschlager et al., 2020) since AR environments can decrease the cognitive and functional load on the user (Billinghurst and Kato, 1999). However, these works are with external displays such as tabletops and large displays (Fig. 8b), which are less free to move (Ens et al., 2021b; Butscher et al., 2018). For this reason, the FIESTA collaborative system (Lee et al., 2021b) is built in an unconstrained VR environment (Fig. 8a), whereby users can move freely.

In terms of interaction channels, collaborative interactions are more based on voice chat, through co-located or remote communication. For example, Embodied Axes (Cordeil et al., 2020) deployed the system to a remote client, but the communication is still through third-party video calling software. In addition to sonic communication, gaze, deictic pointing gestures or placement gestures can also indicate collaborator's attention focus to facilitate communication (Fig. 8b). Moreover, Lee et al. mentioned that the use of avatars and pointers also facilitated collaboration, with deixis allowing participants to work while up close or when far apart (Lee et al., 2021b, Fig. 8a).

5. Application status

Since metaverse applications are still in their infancy, we investigated potential applications that might be the core applications of the future metaverse, such as visual social, virtual medical, virtual city or virtual games, to prospect what the future metaverse might look like.

5.1. *Virtual social*

A number of virtual platforms have been proposed that officially support social networking. These platforms have strong collaborative interaction capabilities, such as Rec Room,[9] VRChat[10] and AltspaceVR.[11] These prototypes are different from each other in the aspects of navigation, spaces, and social mechanics. For example, Rec Room and AltspaceVR use teleportation as the primary mode of navigation. Rec Room supports more embodied modes of friending (through a virtual "handshake") as well as muting (by putting one's hand up as if to say "stop"). Recently, Duan et al. implemented an initial micro metaverse prototype of a university campus, which supports rich virtual social activities (Duan et al., 2021). These platforms have very strong collaborative communication capabilities through high fidelity graphical interface and rich interactive approaches.

5.2. Virtual medical

Metaverse can help medical professionals work faster, cleaner, more virtual, and safer when caring for their patients. For example, AccuVein[12] projects a map of a patient's veins onto the skin. Beyond Metaverse[13] makes innovative extended reality (XR) solutions to improve medical education, training for clinicians, surgical planning, procedure, treatment, and diagnosis. Among these applications, we found that visualization aids to understanding are most used in the medical field and are most closely linked to related scientific visualization research, but users in these applications are usually co-located to explore in AR environments.

5.3. Virtual city

Although there are fewer specialized platforms for virtual cities, more integrated into other areas. We found that the simulation of transport and reconstruction of urban buildings were more prominent in the few applications. For example, MegaWorld[14] is an open platform with avatars, public transport such as subways and buses, etc. Users can choose from any mode of transportation, such as different bus routes, to experience real city life. In this world, handcraft is allowed users to create something by themselves. In addition, like real city life, the core of this application is the vibrant player-driven economics, such as trade and taxes. We can imagine that the future metaverse platform can simulate real government affairs and life business, so as to facilitate the management of cities. The scenes in current applications are very similar to real cities, which means that the high fidelity city reconstruction technology will become an important foundation for the realization of the combination of the metaverse and smart city.

5.4. Virtual games

Virtual games are the closest applications to the concept of a metaverse, which are much more flexible once they are attached to holographic virtual reality environments with futuristic technologies. Bringing the metaverse to the global stage was the listing of Roblox.[15] Compared to other games, Roblox games have their own characters, focus on social needs, which is considered an early form of the metaverse. Other games or apps that most closely resemble the form of the metaverse includes: Decentraland,[16] The Sandbox,[17] Cryptovoxels.[18] Compared with the framework we have proposed for metaverse, we found that virtual games have the positive qualities, high realism, freedom, and high sharing and sociality. However, the visual representation and exploration of the metaverse still have further requirements to be explored, such as complex exploration and perception tasks for single or multiple users, with very high demands for real-time rendering, interaction and visualization.

6. Opportunities for future research

In this section, we discuss the future research opportunities of graphics for visual construction of metaverse, and that of interaction and visualization for user-centered exploration.

6.1. Visual construction of metaverse

Building a more realistic virtual world. As technology continues to innovate and the real world evolves at a rapid pace, how to make the metaverse more realistic will need to be further developed through graphic technology. For example, the realism of the avatar's dress code and fabrics, and the high fidelity of the faces, all require more accurate modeling algorithms. This high realism is also reflected in the creation of animations in the metaverse, where people often see dynamic and interactive scenes and NPCs and avatars, however, everything in the world cannot rely entirely on manual modeling and requires certain automation. By reviewing the literature, we have found that deep learning-based techniques help to improve 3D construction, and thus modeling with high accuracy, automation and interactive VR animation construction is a future challenge and research direction.

Incorporating human creativity. Real-world people can be the inspiration and source of many products in the metaverse. The construction of the metaverse should allow for user input in it, which is then generated with the aid of graphical techniques. In addition to the creation of avatars, many 3D artifacts of a person, such as dwellings, vehicles, etc., could be allowed to be reconstructed in the metaverse. This places a higher demand on image-based or video-based reconstructions taken by portable devices. In addition, some 2D objects including images, videos, artworks, visualizations or human–computer interfaces created by people, can be transferred into metaverse and combined with 3D objects to enable better immersive experiences.

6.2. User interaction in metaverse

Reducing user interaction burden. The variety of user actions depends on the development of sensors and devices. The types of action input captured by various devices can fulfill the requirements of many different types of immersive interaction. As mentioned before, there are various devices designed to support users to navigate, touch and edit things in the metaverse. Users can use their body language, such as gaze, gesture, or head positions, besides, users can use action through sonification and haptic channels. Users can also take action simply by thoughts, that is, using brain signals such as EEG. There are a few limits. Types of user action are limited due to the design mechanism of different developers and are still different from natural interaction action that happens in the real world. Users are required to facilitate certain devices and to remember specific operations designed by different developers, applications and platforms.

---

[9] https://recroom.com.

[10] https://hello.vrchat.com/.

[11] https://altvr.com/.

[12] https://www.accuvein.com/.

[13] https://www.veyondmetaverse.com/.

[14] https://megaworld.io.

[15] https://www.roblox.com/.

[16] https://decentraland.org.

[17] https://www.sandbox.game/en/.

[18] https://www.cryptovoxels.com/.

Feedback with multi-sensory channels. From the visual channel, acoustic channel and haptic channel to olfactory channel and gustatory channel, types of feedback of users have been researched. Various devices try to restore the realism of the virtual environment from all kinds of aspects as much as possible. Improvements are implemented in all aspects to refine the immersive user experience. For example, via the gustatory channel, instead of using chemicals to simulate a taste, other methods such as electronic, thermal and magnetic are applied to reduce the use of physical instances, which always creates aftermarket problems such as the need for refilling. We expect to see the emergence of tighter, all-in-one devices that better combine multiple sensory channels to better immerse users in the metaverse.

6.3. *User awareness in metaverse*

Enriching visualizations in metaverse. Data visualization in the real world has already penetrated all aspects of life and is deeply integrated with different kinds of data to assist people's perception of life, but in the metaverse, visualization is not yet widely used and people's experience of information perception in the metaverse is not yet sufficient. In the future, if people work and study in the metaverse, or even present their work remotely, it is an open question how complex data, such as network structures, trees, 3D scatters can be reasonably perceived and analyzed. Moreover, in order to better experience the metaverse and understand things in it, data visualization should be more deeply integrated into the basic elements of the metaverse, such as scene self-description, storytelling, etc.

Associating visualization with user needs. How to rationally link visualization to user needs is the main challenge to providing a user-centric awareness experience, the main manifestations are when, where and how. First, when the user will need this information may require the user behavioral data (trajectories, clicks, eye movements, etc.) to develop an information-assisted strategy or automatic recommendations with the help of machine learning algorithms. Second, the pre-set arrangement of such visualization and whether its presence interferes with the observation of other things require empirical analysis of the interaction design. Third, the way to interact with the visualization can be controlled by the user or can be done by NPCs. For example, intelligent robots can help users perceive the metaverse environment and the visualizations by offering hints, communicating and interacting with users.

Enhancing collaboration in visualization. One of the advantages for users to engage with the metaverse is that it comes with social properties that facilitate collaborative communication. As the user perceives the virtual world, whose personal insights and observations should be passed on to others. This transmission is sometimes limited to co-located shared observations with voice communication. Other methods of interaction and sensory channels should be explored to aid collaboration. It is worth noting that the avatar plays an important role that allows collaborators to quickly and naturally learn what the user wants to convey, just as in the real world.

7. Conclusions

In this work, we present a framework describing metaverse from the perspectives of graphics, interaction and visualization. We first describe the graphical techniques used to construct visual elements of metaverse (scenes, NPCs and avatars) through a technical pipeline within a taxonomy. And we propose two taxonomies to summarize the research status of interaction and visualization techniques that can support user interaction and awareness of such visual elements in the metaverse. Through our framework, scholars in related fields can readily know how to create visual elements of the metaverse. We investigate the related potential applications for metaverse in the fields of virtual games, virtual social, virtual medical or city), and prospect what the metaverse might look like in the future. Finally, we discuss the research opportunities for future work based on our review. We believe this survey can provide useful insights into the field of visual construction and exploration of the metaverse.

CRediT authorship contribution statement

Yuheng Zhao: Conceptualization, Formal analysis, Writing – original draft. **Jinjing Jiang:** Writing – original draft. **Yi Chen:** Writing – review & editing. **Richen Liu:** Writing – review & editing. **Yalong Yang:** Writing – review & editing. **Xiangyang Xue:** Project administration, Resources. **Siming Chen:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by Shanghai Municipal Science and Technology Major Project (No. 2018SHZDZX01, 2021SHZDZX0103) and ZJLab. This work is also supported by Shanghai Sailing Program No. 21YF1402900, Science and Technology Commission of Shanghai Municipality (Grant No. 21ZR1403300), and supported by Open Research Fund of Beijing Key Laboratory of Big Data Technology for Food Safety (Project No. BTBD-2021KF03), Beijing Technology and Business University and NSFC No. 61972010.

Ethical Approval

This study does not contain any studies with human or animal subjects performed by any of the authors.

References

Bargteil, A.W., Shinar, T., Kry, P.G., 2020. An introduction to physics-based animation. In: SIGGRAPH Asia 2020 Courses. In: SA '20, Association for Computing Machinery, New York, NY, USA.

Batch, A., Cunningham, A., Cordeil, M., Elmqvist, N., Dwyer, T., Thomas, B.H., Marriott, K., 2019. There is no spoon: Evaluating performance, space use, and presence with expert domain users in immersive analytics. IEEE Trans. Vis. Comput. Graphics 26 (1), 536–546.

Batch, A., Patnaik, B., Akazue, M., Elmqvist, N., 2020. Scents and sensibility: Evaluating information olfactation. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–14.

Beck, S., Kunert, A., Kulik, A., Froehlich, B., 2013. Immersive group-to-group telepresence. IEEE Trans. Vis. Comput. Graphics 19 (4), 616–625.

Beer, S., 2015. Virtual museums: an innovative kind of museum survey. In: Proceedings of the 2015 Virtual Reality International Conference. pp. 1–6.

Besançon, L., Ynnerman, A., Keefe, D.F., Yu, L., Isenberg, T., 2021. The state of the art of spatial interfaces for 3d visualization. In: Computer Graphics Forum. Vol. 40. No. 1. Wiley Online Library, pp. 293–326.

Billinghurst, M., Kato, H., 1999. Collaborative mixed reality. In: Proceedings of the First International Symposium on Mixed Reality. pp. 261–284.

Booth, J., Roussos, A., Zafeiriou, S., Ponniah, A., Dunaway, D., 2016. A 3d morphable model learnt from 10,000 faces. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 5543–5552.

Bouzbib, E., Bailly, G., Haliyo, S., Frey, P., 2021. " Can I touch this?": Survey of virtual reality interactions via haptic solutions. arXiv preprint arXiv:2101.11278.

Butscher, S., Hubenschmid, S., Müller, J., Fuchs, J., Reiterer, H., 2018. Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multidimensional Data. Association for Computing Machinery, New York, NY, USA, pp. 1–12.

Cantu, A., Duval, T., Grisvard, O., Coppin, G., 2018. Helovis: A helical visualization for sigint analysis using 3d immersion. In: 2018 IEEE Pacific Visualization Symposium. PacificVis, IEEE, pp. 175–179.

Cordeil, M., Bach, B., Cunningham, A., Montoya, B., Smith, R.T., Thomas, B.H., Dwyer, T., 2020. Embodied axes: Tangible, actuated interaction for 3D augmented reality data spaces. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. In: CHI '20, Association for Computing Machinery, New York, NY, USA, pp. 1–12.

Cordeil, M., Cunningham, A., Dwyer, T., Thomas, B.H., Marriott, K., 2017. Imaxes: Immersive axes as embodied affordances for interactive multivariate data visualisation. In: UIST '17, Association for Computing Machinery, New York, NY, USA, pp. 71–83.

Davis, A., Murphy, J.D., Owens, D., Khazanchi, D., Zigurs, I., 2009. Avatars, people, and virtual worlds: Foundations for research in metaverses. J. Assoc. Inf. Syst. 10 (2), 90.

Dehesa, J., Vidler, A., Lutteroth, C., Padget, J., 2020. Touché: Data-driven interactive sword fighting in virtual reality. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–14.

Deng, T., Zhang, K., Shen, Z.-J.M., 2021. A systematic review of a digital twin city: A new pattern of urban governance toward smart cities. J. Manag. Sci. Eng.

Duan, H., Li, J., Fan, S., Lin, Z., Wu, X., Cai, W., 2021. Metaverse for social good: A university campus prototype. In: Proceedings of the 29th ACM International Conference on Multimedia. Association for Computing Machinery, New York, NY, USA, pp. 153–161.

Ens, B., Bach, B., Cordeil, M., Engelke, U., Serrano, M., Willett, W., Prouzeau, A., Anthes, C., Büschel, W., Dunne, C., et al., 2021a. Grand challenges in immersive analytics. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. pp. 1–17.

Ens, B., Goodwin, S., Prouzeau, A., Anderson, F., Wang, F.Y., Gratzl, S., Lucarelli, Z., Moyle, B., Smiley, J., Dwyer, T., 2021b. Uplift: A tangible and immersive tabletop system for casual collaborative visual analytics. IEEE Trans. Vis. Comput. Graphics 27 (2), 1193–1203.

Ens, B., Hincapié-Ramos, J.D., Irani, P., 2014. Ethereal planes: a design framework for 2D information space in 3D mixed reality environments. In: Proceedings of the 2nd ACM Symposium on Spatial User Interaction. pp. 2–12.

Eynard, R., Pallot, M., Christmann, O., Richir, S., 2015. Impact of verbal communication on user experience in 3D immersive virtual environments. In: 2015 IEEE International Conference on Engineering, Technology and Innovation/International Technology Management Conference. ICE/ITMC, IEEE, pp. 1–8.

Filho, J.A.W., Freitas, C.M., Nedel, L., 2018. VirtualDesk: A Comfortable and efficient immersive information visualization approach. Comput. Graph. Forum.

Fonnet, A., Prie, Y., 2019. Survey of immersive analytics. IEEE Trans. Vis. Comput. Graphics.

Freiknecht, J., Effelsberg, W., 2017. A survey on the procedural generation of virtual worlds. Multimodal Technol. Interact. 1 (4), 27.

Friedman, D., Leeb, R., Guger, C., Steed, A., Pfurtscheller, G., Slater, M., 2007. Navigating virtual reality by thought: What is it like? Presence Teleoperators Virtual Environ. 16 (1), 100–110.

Genay, A.C.S., Lecuyer, A., Hachet, M., 2021. Being an avatar" for real": a survey on virtual embodiment in augmented reality. IEEE Trans. Vis. Comput. Graphics.

Greenwald, S.W., Corning, W., Maes, P., 2017. Multi-user framework for collaboration and co-creation in virtual reality. In: 12th International Conference on Computer Supported Collaborative Learning.

Grubert, J., Itoh, Y., Moser, K., Swan, J.E., 2017. A survey of calibration methods for optical see-through head-mounted displays. IEEE Trans. Vis. Comput. Graphics 24 (9), 2649–2662.

Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., Bennamoun, M., 2021. Deep learning for 3D point clouds: A survey. IEEE Trans. Pattern Anal. Mach. Intell. 43 (12), 4338–4364.

Huang, X., Mei, G., Zhang, J., Abbas, R., 2021. A comprehensive survey on point cloud registration. arXiv preprint arXiv:2103.02690.

Hurter, C., Riche, N.H., Drucker, S.M., Cordeil, M., Alligier, R., Vuillemot, R., 2018. Fiberclay: Sculpting three dimensional trajectories to reveal structural insights. IEEE Trans. Vis. Comput. Graphics 25 (1), 704–714.

Intwala, A.M., Magikar, A., 2016. A review on process of 3D model reconstruction. In: 2016 International Conference on Electrical, Electronics, and Optimization Techniques. ICEEOT, IEEE, pp. 2851–2855.

Ivanov, A., Danyluk, K.T., Willett, W., 2018. Exploration & anthropomorphism in immersive unit visualizations. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. pp. 1–6.

Jankowski, J., Hachet, M., 2013. A survey of interaction techniques for interactive 3D environments. In: Eurographics 2013-STAR.

Johnson, D., Damian, D., Tzanetakis, G., 2019. Osc-xr: A toolkit for extended reality immersive music interfaces. In: Proc. Sound Music Comput. Conf. pp. 202–209.

Joshua, J., 2017. Information bodies: Computational anxiety in neal stephenson's snow crash. Interdiscip. Lit. Stud. 19, 17–47.

Kaplan, O., Yamamoto, G., Yoshitake, Y., Taketomi, T., Sandor, C., Kato, H., 2016. In-situ visualization of pedaling forces on cycling training videos. In: 2016 IEEE International Conference on Systems, Man, and Cybernetics. SMC, IEEE, pp. 000994–000999.

Karunanayaka, K., Johari, N., Hariri, S., Camelia, H., Bielawski, K.S., Cheok, A.D., 2018. New thermal taste actuation technology for future multisensory virtual reality and internet. IEEE Trans. Vis. Comput. Graphics 24 (4), 1496–1505.

Kelly, J.W., Cherep, L.A., Lim, A.F., Doty, T., Gilber, S.B., 2021. Who are virtual reality headset owners? a survey and comparison of headset owners and non-owners. In: 2021 IEEE Virtual Reality and 3D User Interfaces. VR, IEEE, pp. 687–694.

Khaloo, A., Lattanzi, D., 2017. Hierarchical dense structure-from-motion reconstructions for infrastructure condition assessment. J. Comput. Civ. Eng. 31 (1), 04016047.

Kocur, M., Graf, S., Schwind, V., 2020. The impact of missing fingers in virtual reality. In: 26th ACM Symposium on Virtual Reality Software and Technology. In: VRST '20, Association for Computing Machinery, New York, NY, USA.

Kraus, M., Fuchs, J., Sommer, B., Klein, K., Engelke, U., Keim, D., Schreiber, F., 2021. Immersive analytics with abstract 3D visualizations: A survey. In: Computer Graphics Forum. Wiley Online Library.

Latif, S., Tarner, H., Beck, F., 2021. Talking realities: Audio guides in virtual reality visualizations. IEEE Comput. Graph. Appl.

Lau, M., Kuffner, J.J., 2005. Behavior planning for character animation. In: Proceedings of the 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation. pp. 271–280.

Lécuyer, A., Lotte, F., Reilly, R.B., Leeb, R., Hirose, M., Slater, M., 2008. Brain-computer interfaces, virtual reality, and videogames. Computer 41 (10), 66–72.

Lee, L.-H., Braud, T., Zhou, P., Wang, L., Xu, D., Lin, Z., Kumar, A., Bermejo, C., Hui, P., 2021a. All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda. arXiv preprint arXiv:2110.05352.

Lee, B., Hu, X., Cordeil, M., Prouzeau, A., Jenny, B., Dwyer, T., 2021b. Shared surfaces and spaces: Collaborative data visualisation in a co-located immersive environment. IEEE Trans. Vis. Comput. Graphics 27, 1171–1181.

Lee, L.-H., Hui, P., 2018. Interaction methods for smart glasses: A survey. IEEE Access 6, 28712–28732.

Leotta, M.J., Long, C., Jacquet, B., Zins, M., Lipsa, D., Shan, J., Xu, B., Li, Z., Zhang, X., Chang, S.-F., et al., 2019. Urban semantic 3D reconstruction from multiview satellite imagery. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops.

Li, Y., Feng, C., Yu, H., Pang, L., 2021. A survey of physics-based character animation synthesis methods. Animation 3, 41–42.

Lin, T., Singh, R.P., Yang, Y., Nobre, C., Beyer, J., Smith, M.A., Pfister, H., 2021. Towards an understanding of situated AR visualization for basketball free-throw training. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.

Liu, Y.-F., Cho, S., Spencer Jr., B., Fan, J.-S., 2016. Concrete crack assessment using digital image processing and 3D scene reconstruction. J. Comput. Civ. Eng. 30 (1), 04014124.

Liu, L., Habermann, M., Rudnev, V., Sarkar, K., Gu, J., Theobalt, C., 2021a. Neural actor: Neural free-view synthesis of human actors with pose control. arXiv preprint arXiv:2106.02019.

Liu, C., Han, Y., Jiang, R., Yuan, X., 2021b. Advisor: Automatic visualization answer for natural-language question on tabular data. In: 2021 IEEE 14th Pacific Visualization Symposium. PacificVis, IEEE, pp. 11–20.

Liu, L., Hodgins, J., 2017. Learning to Schedule Control Fragments for Physics-Based Characters Using Deep Q-Learning. Vol. 36. No. 4. Association for Computing Machinery, New York, NY, USA.

Liu, R., Wang, H., Zhang, C., Chen, X., Wang, L., Ji, G., Zhao, B., Mao, Z., Yang, D., 2021c. Narrative scientific data visualization in an immersive environment. Bioinformatics.

Llobera, J., Booth, J., Charbonnier, C., 2021. New techniques in interactive character animation. In: ACM SIGGRAPH 2021 Courses. pp. 1–6.

Luo, H., Nagano, K., Kung, H.-W., Xu, Q., Wang, Z., Wei, L., Hu, L., Li, H., 2021. Normalized avatar synthesis using stylegan and perceptual refinement. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11662–11672.

Lyu, Q., Chai, M., Chen, X., Zhou, K., 2020. Real-time hair simulation with neural interpolation. IEEE Trans. Vis. Comput. Graphics (01), 1.

Ma, X., Kang, K., Zhu, R., Wu, H., Zhou, K., 2021. Free-Form Scanning of Non-Planar Appearance with Neural Trace Photography. Vol. 40. No. 4. Association for Computing Machinery, New York, NY, USA.

Ma, Z., Liu, S., 2018. A review of 3D reconstruction techniques in civil engineering and their applications. Adv. Eng. Inf. 37, 163–174.

Mueller, F., Bernard, F., Sotnychenko, O., Mehta, D., Sridhar, S., Casas, D., Theobalt, C., 2018. Ganerated hands for real-time 3D hand tracking from monocular RGB. In: Proceedings of Computer Vision and Pattern Recognition. CVPR.

Murphy, D., 2017. Building a hybrid virtual agent for testing user empathy and arousal in response to avatar (micro-)expressions. In: Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology. In: VRST '17, Association for Computing Machinery, New York, NY, USA.

Navarro, F., Fdez, J., Garzón, M., Roldán, J.J., Barrientos, A., 2017. Integrating 3D reconstruction and virtual reality: A new approach for immersive teleoperation. In: ROBOT.

Nguyen, A., Le, B., 2013. 3D point cloud segmentation: A survey. In: 2013 6th IEEE Conference on Robotics, Automation and Mechatronics. RAM, pp. 225–230.

Onorati, T., Díaz, P., Zarraonandia, T., Aedo, I., 2018. The immersive bubble chart: a semantic and virtual reality visualization for big data. In: The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings. pp. 176–178.

Patnaik, B., Batch, A., Elmqvist, N., 2018. Information olfactation: Harnessing scent to convey data. IEEE Trans. Vis. Comput. Graphics 25 (1), 726–736.

Perez-Gonzalez, J., Luna-Madrigal, F., Piña-Ramirez, O., 2019. Deep learning point cloud registration based on distance features. IEEE Latin Am. Trans. 17 (12), 2053–2060.

Prouzeau, A., Cordeil, M., Robin, C., Ens, B., Thomas, B.H., Dwyer, T., 2019. Scaptics and highlight-planes: Immersive interaction techniques for finding occluded features in 3D scatterplots. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 1–12.

Prouzeau, A., Wang, Y., Ens, B., Willett, W., Dwyer, T., 2020. Corsican twin: Authoring in situ augmented reality visualisations in virtual reality. In: Proceedings of the International Conference on Advanced Visual Interfaces. In: AVI '20, Association for Computing Machinery, New York, NY, USA.

Pu, S., Vosselman, G., et al., 2006. Automatic extraction of building features from terrestrial laser scanning. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 36 (5), 25–27.

Raaen, K., Sørum, H., 2019. Survey of interactions in popular VR experiences.. In: NIK.

Ranasinghe, N., Karunanayaka, K., Cheok, A.D., Fernando, O.N.N., Nii, H., Gopalakrishnakone, P., 2011. Digital taste and smell communication. In: Proceedings of the 6th International Conference on Body Area Networks. pp. 78–84.

Reipschlager, P., Flemisch, T., Dachselt, R., 2020. Personal augmented reality for information visualization on large interactive displays. IEEE Trans. Vis. Comput. Graphics 27 (2), 1182–1192.

Richard, E., Tijou, A., Richard, P., Ferrier, J.-L., 2006. Multi-modal virtual environments for education with haptic and olfactory feedback. Virtual Real. 10 (3), 207–225.

Romero, J., Tzionas, D., Black, M.J., 2017. Embodied hands: Modeling and capturing hands and bodies together. ACM Trans. Graph. 36 (6).

Roth, D., Latoschik, M.E., Vogeley, K., Bente, G., 2015. Hybrid avatar-agent technology–a conceptual step towards mediated "social" virtual reality and its respective challenges. I-Com 14 (2), 107–114.

Roth, D., Waldow, K., Latoschik, M.E., Fuhrmann, A., Bente, G., 2017. Socially immersive avatar-based communication. In: 2017 IEEE Virtual Reality. VR, IEEE, pp. 259–260.

Saragih, J.M., Lucey, S., Cohn, J.F., 2011. Real-time avatar animation from a single image. In: 2011 IEEE International Conference on Automatic Face and Gesture Recognition. FG, IEEE, pp. 117–124.

Satriadi, K.A., Ens, B., Cordeil, M., Czauderna, T., Jenny, B., 2020. Maps around me: 3D multiview layouts in immersive spaces. Proc. ACM Hum. Comput. Interact. 4 (ISS), 1–20.

Shi, M., Aberman, K., Aristidou, A., Komura, T., Lischinski, D., Cohen-Or, D., Chen, B., 2020. Motionet: 3d human motion reconstruction from monocular video with skeleton consistency. ACM Trans. Graph. 40 (1), 1–15.

Smiley, J., Lee, B., Tandon, S., Cordeil, M., Besançon, L., Knibbe, J., Jenny, B., Dwyer, T., 2021. The MADE-axis: A modular actuated device to embody the Axis of a data dimension. Proc. ACM Hum. Comput. Interact. 5 (ISS), 1–23.

Son, H., Kim, C., Kim, C., 2015. Fully automated as-built 3D pipeline extraction method from laser-scanned data based on curvature computation. J. Comput. Civ. Eng. 29 (4), B4014003.

Sorger, J., Arleo, A., Kán, P., Knecht, W., Waldner, M., 2021. Egocentric network exploration for immersive analytics. In: Computer Graphics Forum. Vol. 40. No. 7. Wiley Online Library, pp. 241–252.

Spielberg, S., Silvestri, A., Penn, Z., Cline, E., De Line, D., 2018. Ready Player One. Warner Bros USA.

Stotko, P., Krumpen, S., Hullin, M.B., Weinmann, M., Klein, R., 2019. SLAMCast: Large-scale, real-time 3D reconstruction and streaming for immersive multi-client live telepresence. IEEE Trans. Vis. Comput. Graphics 25 (5), 2102–2112.

Tang, Y.M., Ho, H.L., 2020. 3D modeling and computer graphics in virtual reality. In: Mixed Reality and Three-Dimensional Computer Graphics. IntechOpen.

Usher, W., Klacansky, P., Federer, F., Bremer, P.-T., Knoll, A., Yarch, J., Angelucci, A., Pascucci, V., 2018. A virtual reality visualization tool for neuron tracing. IEEE Trans. Vis. Comput. Graphics 24 (1), 994–1003.

Vogel, D., Lubos, P., Steinicke, F., 2018. AnimationVR-interactive controller-based animating in virtual reality. In: 2018 IEEE 1st Workshop on Animation in Virtual and Augmented Environments. ANIVAE, IEEE, pp. 1–6.

Wagner, J., Stuerzlinger, W., Nedel, L., 2021. The effect of exploration mode and frame of reference in immersive analytics. IEEE Trans. Vis. Comput. Graphics.

Wang, Y., Solomon, J.M., 2019. Deep closest point: Learning representations for point cloud registration. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. ICCV.

Warpefelt, H., Verhagen, H., 2015. Towards an updated typology of non-player character roles. In: Proceedings of the International Conference on Game and Entertainment Technologies. pp. 1–9.

Wei, X., Yin, L., Zhu, Z., Ji, Q., 2004. Avatar-mediated face tracking and lip reading for human computer interaction. In: Proceedings of the 12th Annual ACM International Conference on Multimedia. In: MULTIMEDIA '04, Association for Computing Machinery, New York, NY, USA, pp. 500–503.

Weissker, T., Bimberg, P., Froehlich, B., 2020. Getting there together: Group navigation in distributed virtual environments. IEEE Trans. Vis. Comput. Graphics 26 (5), 1860–1870.

White, S., Feiner, S., 2009. Sitelens: Situated visualization techniques for urban site visits. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1117–1120.

Wu, M., Wang, Y., Hu, Q., Yu, J., 2020. Multi-view neural human rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1682–1691.

Yang, Y., Cordeil, M., Beyer, J., Dwyer, T., Marriott, K., Pfister, H., 2021. Embodied navigation in immersive abstract data visualization: Is overview+detail or zooming better for 3D scatterplots? IEEE Trans. Vis. Comput. Graphics 27 (2), 1214–1224.

Yang, Y., Dwyer, T., Jenny, B., Marriott, K., Cordeil, M., Chen, H., 2018a. Origin-destination flow maps in immersive environments. IEEE Trans. Vis. Comput. Graphics 25 (1), 693–703.

Yang, Y., Dwyer, T., Marriott, K., Jenny, B., Goodwin, S., 2020. Tilt map: Interactive transitions between choropleth map, prism map and bar chart in immersive environments. IEEE Trans. Vis. Comput. Graphics.

Yang, Y., Jenny, B., Dwyer, T., Marriott, K., Chen, H., Cordeil, M., 2018b. Maps and globes in virtual reality. Comput. Graph. Forum 37 (3), 427–438.