

CHAPTER

11

INTEGRATION AND FUTURE DIRECTIONS

Fazliaty Edora Fadzli, Ajune Wanis Ismail, and Mohd Shafry Mohd Rahim

11.1 INTRODUCTION

In the constantly changing realm of technology, the incorporation of extended reality (XR) has risen as an innovative paradigm, erasing the distinctions between the tangible and virtual realms. XR encompasses a range of immersive technologies, spanning from virtual reality (VR), augmented reality (AR), and mixed reality (MR), all of which seek to augment our sensory experiences in unprecedented ways (Memmesheimer & Ebert, 2022). Figure 11.1 describes the relationships of technology within the umbrella of XR.

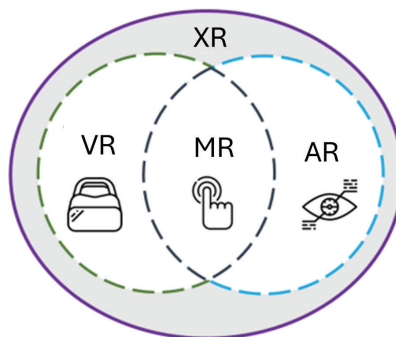


Figure 11.1 The relationship structure for extended reality technology

Virtual reality (VR) allows users to be fully immersed in an imaginary or virtual world, while AR facilitates interactions with virtual objects situated within the real world. On the other hand, MR integrates both AR and VR technologies to generate new environments and visualizations, enabling the coexistence of physical and digital objects (real objects integrated into virtual environments (VE), virtual objects integrated into the real world, or even virtual objects within virtual worlds). This allows for seamless interaction between users, the virtual objects, and the surrounding space. At the forefront of this transformation lies the realm of hand gesture interaction. Hand gestures represent one of the most natural, common, and expressive forms of body language used to convey attitudes and emotions in social interactions. This mode of interaction leverages the innate dexterity and expressiveness of human hands to navigate and manipulate digital environments. By translating real-world gestures into meaningful digital actions, hand gesture interaction holds the potential to redefine human engagement with XR. It bridges the gap between the tactile familiarity of the physical world and the limitless possibilities of the digital realm, promising a more intuitive and seamless user experience.

11.2 EXPANDING GESTURE INTERACTION

Hand gesture interaction plays a pivotal role in making these experiences more intuitive and engaging. To appreciate the current state and future potential of hand gesture interaction in XR, it is essential to trace its evolution. Initially, early VR and AR systems relied on controllers or other input devices. Over time, advances in computer vision and sensor technologies have made hand tracking and gesture recognition feasible, revolutionizing the way users interact with digital content. Within the realm of XR technologies, a diverse array of interaction methods is available, spanning from touch-based interfaces and external controllers to modalities rooted in gaze, speech (Aladin et al., 2020), or mid-air gestures. When designing an MR or VR application, the choice of interaction technique is typically made with

careful consideration of the particular task at hand and the limitations imposed by the underlying technology (Memmesheimer & Ebert, 2022).

Hand gesture recognition forms the backbone of hand gesture interaction in XR. It involves capturing the movements and positions of a user's hands in real-time. Depth-sensing cameras, infrared sensors, and machine learning algorithms are commonly used to achieve accurate hand tracking. Hand gestures serve as a form of non-verbal communication, encompassing both straightforward actions (like pointing at objects) and more intricate expressions of feelings or interaction with others. Instead of relying on specialized devices like Leap Motion, Microsoft Kinect, or other red-green-blue-depth (RGB-D) cameras for their interaction system, they explored the implementation of a deep learning method, making the system accessible to a wider audience. These technologies enable XR systems to interpret hand gestures and translate them into meaningful actions within the virtual environment.

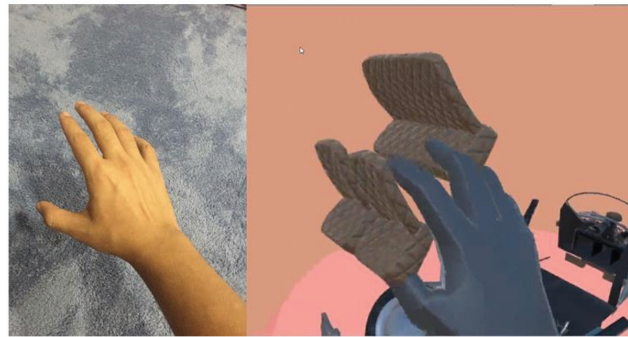


Figure 11.2 User performing grasping interaction (Source: Aladin et al., 2020)

In XR environments, employing hand gesture interaction presents a significant challenge due to the intricate and diverse nature of human hand movements (Sabbarese et al., 2023). In real-life scenarios, gestures are often executed continuously, making it difficult to accurately discern the exact initiation and termination points of a given gesture. The wide range of hand forms and movements further complicates the