The Evolution, Evaluation, and Case Study Analysis of Interactive Paradigms in Virtual Reality Games

Tingxuan Tan

College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, 310023, China
tantingxuan@zjut.edu.cn

Abstract. The fast development of VR hardware has created a gap with relatively slow progress in interaction design, which restricts the improvement of immersive experiences and brings troubles to standardized design and evaluation work. Based on existing literature and typical game cases, the development route of interaction paradigms in VR games is sorted out and a set of analytical and evaluation frameworks is built. A three-stage evolution model and a six-dimensional evaluation model are put forward, and three key paradigms are summarized, including metaphor mapping, compensatory design and intention understanding. The transformation of paradigms comes from the continuous improvement of the deficiencies left by previous design schemes, and intelligent multimodal fusion provides an effective way to upgrade experience quality. The role of interaction has changed from a simple functional tool to a core part that shapes user experience, which is reflected in the ontological change of interaction, embodied narrative and embodied social performance. The models and conclusions in this study can provide stable theoretical support and practical operation guidance for the design and evaluation of natural and intelligent VR interaction systems.

Keywords: Virtual reality games; Human-computer interaction; Natural interaction; Interaction paradigm; Evaluation model

1 Introduction

The core appeal of VR technology lies in building a strong sense of presence for users, and this goal cannot be realized without natural and direct interaction between people and virtual scenes [1]. In recent years, the continuous upgrading of VR hardware perception ability has laid a technical foundation for the landing of natural interaction interfaces [2]. However, the development speed of software interaction paradigms still lags behind the performance improvement of hardware, which has become a prominent bottleneck in the field of interaction design [3]. A large number of VR games still follow the interaction logic of traditional handles, which makes it impossible to give full play to the immersion value brought by body movements [4]. This mismatch between software and hardware leads to the lack of unified design norms and evaluation standards in the industry, which increases the cost and uncertainty in the exploration of

G. E. Atalay and A. Wang (eds.), *Proceedings of the 2026 3rd International Conference on Public Relations and Media Communication (PRMC 2026)*, Advances in Social Science, Education and Humanities Research 1032,
https://doi.org/10.2991/978-2-38476-609-3_72

new interaction forms [5]. The evolutionary logic of VR game interaction from a simple control tool to body-based expression is sorted out through a complete analytical framework. The development process is divided into different stages, the mainstream paradigms in each stage are extracted, and an evaluation system is built to realize the comparison between different paradigms and the prediction of future trends.

2 Literature Review and Theoretical Framework

Key concepts involved in the research are clearly defined to ensure the accuracy and consistency of the analysis process. Natural interaction mainly uses gestures, voice, gaze and other instinctive behaviors to reduce the user's learning cost and improve the fluency and immersion of operation [6]. Interaction paradigm is a reusable design framework that can solve the contradictions between people, machines and environments under specific technical conditions [7]. Embodied interaction attaches importance to the core position of the physical body, movement and sensory perception in the process of interaction and meaning construction. A combination framework including a three-stage evolution model and a six-dimensional evaluation model is built to analyze the development of natural VR interaction, and these two models run through the whole analysis process as core tools.

2.1 Analytical Lens for Evolution Path: Three-Stage Model

According to the maturity of technology and mainstream design ideas, the development of VR interaction is divided into three stages. The device-dependent period takes metaphor mapping as the core, which connects limited physical input with rich virtual operations by simulating real-life behaviors to form an intuitive cognitive path. The limb perception period takes compensatory design as the dominant scheme, which mainly solves the experience problems caused by immature sensing technology such as jitter and occlusion. The full-body integration period focuses on intention understanding, and uses multi-channel sensing and AI context reasoning to narrow the semantic distance between low-level signal input and high-level interaction intention. This model provides a complete historical and logical perspective for the analysis of typical cases and design strategies in each stage.

2.2 Analytical Lens for Design Solutions: Six-Dimensional Evaluation Model

A six-dimensional evaluation model is used to realize the horizontal comparison of the performance of different interaction paradigms in real application scenarios. The six dimensions include immersion and naturalness, learning and cognitive cost, task adaptability and precision, accessibility and comfort, support for narrative and role-playing, and richness of social expression. This model can be used to evaluate mainstream interaction methods and extract effective design strategies. In the following content, the two models are used to analyze and compare Robo Recall, Cubism and Drakheir, so as to explain the connotation, implementation strategy and evolution motivation of metaphor mapping, compensatory design and intention understanding.

3 Core Design Paradigms and Deconstruction of Typical Cases

3.1 Robo Recall: Functional Gamification Design under the Metaphor Mapping Paradigm

Robo Recall is a representative work of the metaphor mapping paradigm in the device-dependent period, and its design core is to make up for the limitation of single controller input. The paradigm reduces the user's cognitive pressure by associating real experience with virtual operations. The parabolic teleportation in the game converts abstract movement into continuous spatial visual operation, which has a significant effect on relieving VR motion sickness. The game also adopts the design idea of functional gamification. The repeated loading action is designed as a performance action of reaching back to grasp the ammunition clip, turning the mechanical operation into rich audio-visual feedback and enhancing the user's emotional participation. The spatial anchor design such as the back holster helps to strengthen the user's spatial cognition and operation memory. Although this paradigm can realize basic functional requirements, it will obviously reduce the naturalness of actual operation.

3.2 Cubism: Systematic Compensation for Technical Shortcomings under the Compensatory Design Paradigm

Cubism presents the application logic of compensatory design in the limb perception period, and the design focus is on adjusting user expectations and making up for the defects of hand tracking such as jitter, virtual hand penetration and lack of tactile feedback [8]. The work does not rely solely on the improvement of hardware performance, but uses advanced interaction and visual design to offset technical limitations. Common strategies include simplifying fine pinching action into contact-based grabbing to realize input compensation, using translucent ghost hand to help users form a cognitive model that adapts to penetration to complete visual compensation, and strengthening multi-sensory feedback through object highlighting and clear sound effects to achieve system compensation. These designs prioritize the stability and reliability of direct natural interaction, even though they are still limited by the level of sensing technology at that time [9].

3.3 Drakheir: Multimodal Fusion Interaction under the Intention Understanding Paradigm

Drakheir represents the frontier exploration direction of the full-body integration period, and the core of its paradigm is to infer high-level operation intention from low-level multi-channel input signals [10]. The system does not rely on strict posture matching, but uses fuzzy judgment based on motion trend and context information, and integrates gesture, voice and gaze input [11]. The fault-tolerant gesture recognition based on motion trend makes the casting action smooth and reduces the difficulty of getting started. The multi-level imprint gesture system with basic vocabulary, combined grammar and usage rules turns interaction into the core content of gameplay and narrative, making operation itself become a part of experience rather than a simple control method. The design of gesture combined with voice casting points to the future development direction of context-aware, adaptive and seamless intelligent interaction.

3.4 Summary of Evolutionary Dynamics

The evolution of VR interaction paradigms is driven by the continuous improvement of unmet experience needs in the previous stage. Metaphor mapping can realize basic functions but lacks natural expression. Compensatory design improves naturalness but is limited by technical conditions. Intention understanding breaks through constraints and moves towards intelligent and context-aware interaction. Each paradigm upgrade aims at the capability defects of the previous scheme and carries out targeted optimization.

4 Comprehensive Evaluation and Design Strategies for Mainstream Interaction Methods

4.1 Evaluation Based on the Six-Dimensional Model

The six-dimensional evaluation model shows obvious performance differences among mainstream interaction methods. Hand tracking performs well in immersion, naturalness and social expression richness [12]. Eye tracking has obvious advantages in reducing learning and cognitive costs, but it is easy to cause visual fatigue and its accuracy is affected by environmental light [13][14]. Voice interaction performs well in immersion, naturalness, task adaptability and precision [15]. Intelligent multimodal fusion interaction has the most balanced ability in all dimensions, and performs prominently in immersion, naturalness, adaptability, precision and narrative support, showing a strong ability to cope with complex interaction needs [16].

4.2 Core Challenges and Response Strategy Analysis

Four common design strategies are summarized from case studies to deal with common technical limitations, including scene transformation, logical fault tolerance, multi-sensory substitution and high-level tactile simulation. Each mainstream interaction method has corresponding solutions for its core problems. Hand tracking may bring physical fatigue and insufficient precision in fine operation [12], which can be improved by transforming continuous fine operation into discrete state judgment as in Cubism, or adopting motion trend-based fault-tolerant recognition as in Drakheir to enhance fluency and stability at the expense of part of posture accuracy. Eye tracking may lead to visual fatigue, and its accuracy is affected by personal differences and light environment [14], so it is suitable for use in key scenarios with high concentration, and can be combined with gestures and head movements to verify intentions and improve reliability [13]. Voice interaction is easily disturbed by noise and is not suitable for long and complex instructions [15], so it needs to be used with gestures and gaze. In Drakheir, voice is used to trigger macro instructions and enrich casting effects, while gestures are responsible for aiming and selection to maximize advantages. Multimodal fusion can fundamentally solve the defects of a single mode [16]. Through intelligent scheduling and context awareness, the system can dynamically select the best input combination for different task scenarios, so as to optimize the overall interaction experience [17].

5 Discussion: Deep Impact of Paradigm Shift and Future Directions

5.1 From Tools to Language: The Evolution of Interactive Ontology

The evolution of VR interaction paradigms has changed the role of interaction from a functional auxiliary tool to a core language that defines virtual experience. In Drakheir, the gesture system integrates operation and content, making interaction the core of gameplay. In Robo Recall, the performative loading design realizes the narrative experience based on body movements. Social interaction also tends to be embodied, relying more on body language and eye contact [18]. These changes mark a fundamental shift in VR design logic, from designing gameplay to serve stories to building experience with the body as the core.

5.2 Future Research Directions

Future research can focus on three key directions. The first is to build an intelligent multimodal fusion model with context awareness to realize seamless and adaptive interaction. The second is to explore AI-driven personalized interaction and use machine learning to adjust parameters for individual users. The third is to promote cross-platform standardization and explore emerging technologies including high-precision tactile equipment and electromyography interfaces [19].

6 Conclusion

Through the analysis of typical VR game cases, the evolutionary logic of interaction paradigms is clarified, which has gone through three stages: metaphor mapping, compensatory design and intention understanding. The core driving force of evolution is to continuously make up for the capability defects of the previous paradigm and move towards body-centered intelligent multimodal fusion. The three-stage evolution model and six-dimensional evaluation model are put forward, and the core design concept is summarized: technical defects should be actively compensated through advanced design strategies rather than passively accepted. The evolution turns VR interaction from a functional tool into a core part of experience, marking the arrival of body-centered experience design. Follow-up research can focus on intelligent multimodal fusion, AI personalized interaction and cross-platform standardization, so as to promote VR interaction to be more natural, intelligent and humanized.

References

1. Geniş-Ergün B. and Şahiner R.: Embodiment in virtual reality and augmented reality games: An investigation on the user interface haptic controllers. *Journal of Soft Computing and Artificial Intelligence* 4(1), 1-13 (2023)
2. Lei Y., Deng Y., Dong L., et al.: A novel sensor fusion approach for precise hand tracking in virtual reality-based human–computer interaction. *Biomimetics* 8(3), 326 (2023)

3. Wu Y., Wang Y., Lou X., et al.: An empirical practice of design and evaluation of freehand interaction gestures in virtual reality. *Multimedia Tools and Applications* 82(52), 39821-39842 (2023)
4. Capece N., Erra U., and Gruoso M.: Evaluation of virtual reality interaction techniques: the case of 3D graph. *arXiv:2302.05660* (2023)
5. Wang Z. M., Rao M. H., and Ye S. H.: Towards spatial computing: recent advances in multimodal natural interaction for extended reality headsets. *Frontiers of Computer Science* 19(12), 1912708 (2025)
6. Zhang Y. X.: Interactive design of online teaching platform based on virtual reality technology. *Journal of Modern Educational Technology* 33(10), 105-111 (2023)
7. Bowman D. A., Kruijff E., Laviola J. J., et al.: *3D user interfaces: theory and practice*. Boston: Addison-Wesley (2004)
8. Mahdi M. B., Guillou E., and Meyer A.: Pure physics-based hand interaction in VR. In *Proceedings of the 19th International Conference on Computer Vision Theory and Applications*, SCITEPRESS, 366-373 (2024)
9. Abdelkarem D., Di L., and Aves P.: Comparison of hand tracking-based and controller-based interaction in a consumer virtual reality game. *Universal Access in the Information Society* 24(3), 1123-1138 (2025)
10. Wang Z., Gu C., and Lu F.: SIAgent: spatial interaction agent via LLM-powered eye-hand motion intent understanding in VR. *arXiv preprint* (2026)
11. Cao L., Zhang H., and Peng C.: Real-time multimodal interaction in virtual reality - a case study with a large virtual interface. *Multimedia Tools and Applications* 82, 12345-12367 (2023)
12. Tseng W.-J., Huron S., and Lecolinet E.: FingerMapper: mapping finger motions onto virtual arms to enable safe virtual reality interaction in confined spaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, New York: ACM, 1-15 (2023)
13. Joos L., Klein K., and Keim D. A.: Exploring trajectory data in augmented reality: a comparative study of interaction modalities. *IEEE Transactions on Visualization and Computer Graphics* 29(5), 2345-2358 (2023)
14. Sidenmark L. and Lundström A.: Gaze behaviour on interacted objects during hand interaction in virtual reality for eye tracking calibration. In *Proceedings of the 2019 ACM Symposium on Eye Tracking Research & Applications*, New York: ACM, 46 (2019)
15. Bolt R. A.: "Put-that-there": voice and gesture at the graphics interface. In *Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques*, New York: ACM, 262-270 (1980)
16. Tao J. H., Wu Y. C., Yu C., et al.: A survey on multimodal human-computer interaction. *Journal of Image and Graphics* 27(6), 1956-1987 (2022)
17. Wagner J., Stuerzlinger W., and Nedel L.: Comparing and combining virtual hand and virtual ray pointer interactions for data manipulation in immersive analytics. *IEEE Transactions on Visualization and Computer Graphics* 27(5), 2513-2523 (2021)
18. Hu Y., Panchanadikar P., and Freeman G.: Beyond traditional gaming: understanding how social VR streaming creates new forms of play. In *Proceedings of the 2025 CHI PLAY Conference on Human-Computer Interaction*, New York: ACM, 1-13 (2025)
19. Townsend L. L., Rasch J., and Grech A.: A tree's perspective: enhancing nature connectedness through transitional and multisensory virtual reality experiences. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*, New York: ACM, 1-13 (2026)