Design of "Virtual Fusion"Cooking Experience Based on XR: Transplantation and Reconstruction of Game Mechanism in Real Skill Training

Ying Wan

Institute of Esports, Communication University of China, Nanjing, China
WanYing20250822@outlook.com

Abstract. At present, there are still some problems in the traditional cooking teaching mode, such as the boring explanation of theoretical knowledge, the non intuitive guidance of practical operation, and the serious disconnection between knowledge teaching and hands-on operation, which restrict the efficiency and experience of cooking learning. However, with the development of extended reality (XR) technology, the digital transformation of cooking teaching has ushered in a new opportunity. Based on the design concept of game, this paper proposes a new "virtual fusion" cooking experience system to solve the problem of the gap between knowledge transfer and actual operation in traditional cooking teaching. Through the three core modules of spatial perception, virtual real mapping and dynamic feedback, the system transforms abstract cooking knowledge into specific spatial action instructions. With the immersion characteristics of XR technology, the system transforms abstract cooking knowledge such as temperature control, food material processing and seasoning proportion into intuitive and landing spatial action instructions, so that beginners can master cooking skills in interesting interaction. Finally, this paper discusses the technical implementation path of the system and Its Enlightenment for the development of intelligent kitchen in the future.

Keywords: Expand reality, Cooking teaching, Game design, Fusion of emptiness and reality

1. Introduction

Cooking is a comprehensive skill training activity. For a long time, it has been taught by relying on traditional picture and text recipes and video tutorials, but this teaching method has the problems of heavy cognitive burden and poor learning effect [1]. In recent years, with the development of augmented reality (AR) and virtual reality (VR) technology, the market scale of intelligent kitchen is expected to reach US $43.8 billion by 2029, with a compound annual growth rate of 19.2% [2]. Extended reality (XR) technology has brought a new change to cooking teaching, projecting virtual information into the real world, so that people can "see what they get", so as to obtain an immersive learning experience [3]. However, there is still a clear gap in the relevant research. Most software applications only stay at the level of menu display and video teaching, lacking accurate guidance and real-time feedback on cooking movements. The deep integration of game mechanism and real

cooking skill training is insufficient, and a systematic and landing teaching assistance framework has not been formed. Therefore, based on ar Augmented Reality Theory, this paper designs an AR assisted cooking system framework to explore the application value of Gamification method in practical skill training, aiming to make up for the deficiencies of existing research in interaction design, skill training standardization and learning incentive mechanism, and achieve the research goal of reducing cooking learning threshold and improving skill mastery efficiency by optimizing the virtual reality fusion experience, improving learning interest and training effectiveness, so as to provide theoretical support and practical reference for the development of intelligent cooking teaching and intelligent kitchen scene.

2. Theoretical basis and related work

2.1. Application of traditional theory in skill training

According to the traditional theory of cooking, physical feeling and cognition are inseparable, and the process of skill learning is a circular process of "perception action feedback" [4]. However, the learning of cooking skills requires the ability to perceive many aspects, such as knife work, temperature, and the feeling of food materials [5]. Traditional recipes simplify complex actions into words, but this leads to a gap between knowledge and practice [6]. AR technology uses spatial and visual guidance to display, which can turn abstract cooking knowledge into specific guidance, so that learners can get timely feedback in actual operation [7]. This is the way provided by virtual reality in the new era, which is conducive to the learning of skills and the formation of muscle memory.

2.2. Development status of XR technology in the field of education and training

XR technology has developed rapidly in education and training. The scale of AR education market worldwide is expected to reach US $5.78 billion by 2025 [8]. Mobile augmented reality (mobile AR) and web-based AR technology bring great convenience to educational applications [9]. In culinary education, the current research mainly focuses on the basic visual guidance and step display [10,11]. Eye tracking and voice control technology make interaction more natural and convenient, while non immersive VR system has good teaching effect for special groups [12]. However, at present, most researchers are inclined to study how to assist pathological personnel in cooking, so there is a lack of game design and in-depth skill evaluation of cooking skills, which affects its effect in learning complex cooking skills. The trend analysis of AR technology for agricultural food applications in recent years shows that real-time image tracking and intelligent feedback are important factors to improve user experience [13,14].

2.3. Theoretical framework and design principles of gamification mechanism

The Gamification mechanism can greatly improve learners' enthusiasm and persistence by setting goals, progress feedback, achievement rewards and other ways. In skill training, reasonable Gamification design should focus on three aspects: timely feedback, gradually increasing difficulty, and social interaction [15]. The culinary Gamification system needs to be both interesting and practical, but it cannot affect the real skill learning too much. Successful cooking VR game applications generally have a hierarchical incentive system, real-time feedback on operation (such as the light effect when cutting vegetables), progress display on steps (such as the completion of recipes), and achievement unlocking on skills (such as the improvement of knife level). Such a

multi-level incentive mode can maintain the enthusiasm of learners at all stages and is conducive to the progress from novice to expert.

3. Design of virtual fusion cooking system

3.1. Overall system architecture and core modules

As illustrated in Figure 1, the virtual fusion cooking system uses a three-tier closed-loop architecture model, forming a technology chain from perceptual input to feedback output. The system is divided into three parts: perception layer, processing layer and interaction layer, so as to realize the technical closed loop of "perception cognition feedback". The perception layer uses a variety of sensors to scan the kitchen in an all-round way, the processing layer uses artificial intelligence to transform traditional recipes into executable operation steps, and the interaction layer uses ar devices to show the operation process to users and give timely feedback. The whole system is designed with the idea of modularization, and each subsystem communicates and cooperates through standard interfaces to facilitate the expansion and upgrading of the system.

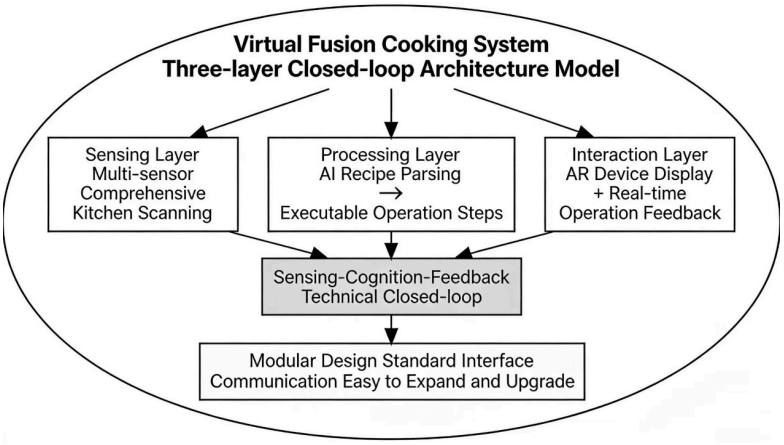

Figure 1. Virtual fusion cooking system (picture credit: original)

3.2. Spatial awareness and environment modeling mechanism

The spatial awareness module is based on Slam (synchronous positioning and map building) technology and deep learning visual algorithm to complete the accurate 3D modeling and tracking of the kitchen environment. The system uses rgb-d camera to obtain the environmental depth information, and coordinates with IMU sensor data for spatial coordinate correction to generate a high-precision point cloud model of kitchen space. The environment modeling method uses semantic segmentation technology to automatically detect and classify kitchen items, and generates a three-dimensional semantic map of important objects such as stove, chopping board, kitchen utensils, food materials and so on. The system uses the constantly updated vision inertial fusion positioning technology to track the position of users' hands and operating tools in real time, so as to ensure that the virtual content matches the real world well. In order to adapt to the changes of the kitchen environment, the system adds the adaptive relocation function, which can still maintain a good spatial positioning effect in the case of light changes, object movement and so on, laying a good foundation for the combination of virtual and real interaction.

3.3. Virtual real mapping and dynamic feedback system

The virtual reality mapping module turns traditional paper recipes into executable spatial actions, that is, turning abstract knowledge into concrete actions. The system uses natural language processing technology to analyze the menu, find out the operation object, tool requirements, action standards and other information, and transform it into visual guide elements in three-dimensional space. The dynamic feedback system uses computer vision and machine learning methods to track the user's actions and changes in food ingredients, and gives timely operation suggestions and error correction. The system can detect whether the direction of cutting vegetables is correct, whether the heat is appropriate, whether the speed of stir frying is appropriate and other details, and visually mark red, sound and other ways to remind, so that the user's behavior conforms to the regulations. The feedback mechanism is a step-by-step guidance process, which changes the prompt strength according to the user's proficiency, and makes it transition from frequent prompt to autonomous operation.

3.4. Game based interactive interface design

The Gamification interactive interface adopts the layered design idea, which is composed of the spatial guidance layer, the information Kanban layer and the incentive feedback layer, bringing the player an immersive feeling. The spatial guidance layer directly displays the operation prompts on the target (Figure 2), directly superimposes the key cooking instructions and dynamic prompts on the real Kitchenware and ingredients, and realizes the intuitive and clear "WYSIWYG" guidance with the help of visual means such as dynamic light and shadow annotation, size scale prompt, operation route preview, arrow guidance, which can effectively reduce the user's understanding cost and action deviation in complex operations. The information Kanban layer uses a translucent suspension interface (Figure 3) to display key information such as cooking progress bar, step-by-step operation list, real-time status monitoring, temperature and time reminders, and supports non-contact interactive methods such as gesture hover, voice command, space touch, etc. to ensure the efficiency of information acquisition without interfering with the actual cooking action. The incentive feedback layer uses the idea of game design (Figure 4) for reference to continuously stimulate the learning motivation and operation enthusiasm of users and strengthen the interest and sustainability of skill training through multiple forms such as achievement badge unlocking, skill rating, real-time operation scoring, special effect dynamic effect completion, task progress reward, etc. The interface design conforms to the principle of cognitive load minimization, reasonably distributes visual attention, and makes virtual information and cooking operations complement each other. At the same time, the system also provides personalized settings. Users can change the interface style and interaction mode according to their preferences, so as to improve the ease of use of the system and user satisfaction.

Figure 2. Spatial guidance layer (picture credit: original)

Figure 3. Information Kanban layer (picture credit: original)

Figure 4. Incentive feedbake layer (picture credit: original)

4. Technology implementation and verification

Multimodal interaction system combines the technologies of vision, hearing and touch, which enables users to communicate with computers in a more natural and convenient way. The visual interaction part uses gesture recognition and eye tracking technology to complete the functions of clicking, sliding, zooming and so on. These functions are based on convolutional neural network to realize the real-time detection and understanding of gestures, so as to ensure that the system can perceive and correct user operations in real time in the dynamic cooking scene. Auditory interaction uses natural language processing and speech synthesis technology, which can realize voice command control, intelligent question and answer and other functions. Users can quickly switch recipe steps and query cooking points through voice, and also provide real-time operation guidance through intelligent question and answer, further reducing the threshold of learning and operation. The whole system uses the method of multimodal fusion, which can automatically determine what the user wants to do, so that different interaction modes can be freely switched and used together, so as to realize the seamless switching and collaborative cooperation of gesture, eye movement, voice and tactile interaction, and greatly improve the interaction flexibility and scene adaptability. The modular design idea is adopted in technology, and each interactive mode communicates through a common event bus to ensure the unity and reliability of the system.

5. Conclusion

The virtual fusion cooking experience system developed in this study based on XR technology has realized the transformation from traditional graphic teaching to immersive spatial interaction. Using the three-tier closed-loop structure and the game based interactive mode, the system can better solve the problem of the separation of knowledge transfer and actual operation in the traditional cooking teaching, and gives a new idea on the learning of life skills. However, the current system still faces several challenges in the actual landing and large-scale application. Multimodal perception in complex kitchen environment is vulnerable to light, occlusion and noise interference, and the recognition stability needs to be further improved. The wearing comfort and long-term endurance of XR devices are still difficult to fully meet the needs of real cooking scenes. The personalized

interaction model is highly dependent on user data, and has some privacy and computational overhead problems. At the same time, there are still errors in the accurate matching of virtual information with real Kitchenware and operation actions. With the development of 5g communication, edge computing and artificial intelligence, the future virtual fusion cooking system will make significant progress in the following aspects: the system will use more accurate biosensors to provide personalized cooking guidance according to users' physiological conditions. Combine blockchain technology to create a decentralized cooking knowledge sharing platform. The application of digital twin technology makes the combination of virtual kitchen and smart home ecosystem. The development of these technologies will promote the intelligent kitchen to be more intelligent, personalized and ecological, and lay a good foundation for building a smart living space in the future.

References

[1] Califano, G., Zhang, T., & amp; Spence, C. (2026). When and Why Do Users Trust AI in the Kitchen? A Hybrid Modelling Approach to the Adoption of AI-Assisted Cooking. International Journal of Human–Computer Interaction, 42(1), 131–143. https: //doi.org/10.1080/10447318.2025.2505154

[2] Geng, G., et al. (2026). Engaging young minds: How smart augmented reality glasses transform learning experiences in AR smart glasses immersive virtual learning environments (AR smart glasses IVLE). Journal of Science Education and Technology, 1–20. https: //doi.org/10.1007/s10956-026-10299-4

[3] Latif, W. B., Yasin, I. M. Y., Rahaman, M. A., Forid, M. S., Islam, M. N., & Hossain, M. Z. (2024). Impact of augmented reality (AR) and virtual reality (VR) on interactive learning systems. Pac Journal of Advanced Engineering and Innovation, 1(1), 23–32. https: //doi.org/10.70818/pjaei.2024.v01i01.016

[4] Alice, I. G. C., Cristina, V. B., & Mariano, R. A. (2019). A virtual versus an augmented reality cooking task based-tools: A behavioral and physiological study on the assessment of executive functions. Frontiers in Psychology, 10, 2529. https: //doi.org/10.3389/fpsyg.2019.02529

[5] Cañada, Y., Torres, C. S., Martinez, A. J., et al. (2024). Characterization and assessment of executive functions through a virtual cooking task in euthymic patients with bipolar disorder. Journal of Psychiatric Research, 178, 349–358. https: //doi.org/10.1016/j.jpsychires.2024.08.007

[6] Yang, Y., Liu, M., Lyu, M., et al. (2025). Does healthy cooking lead to healthy choices? Exploring the impact of a virtual reality cooking game on healthy eating decisions. Games for Health Journal. https: //doi.org/10.1177/2161783X251378648

[7] JakuJakubow, L., Bouck, C. E., Norwine, L., et al. (2025). Enhancing independence: Non-immersive virtual reality for teaching cooking skills to high school students with intellectual disability. Journal of Special Education Technology, 40(2), 252–265. https: //doi.org/10.1177/01626434241277190

[8] Qiao, X., Ren, P., Dustdar, S., Liu, L., Ma, H., & Chen, J. (2019). Web AR: A promising future for mobile augmented reality—State of the art, challenges, and insights. Proceedings of the IEEE, 107(4), 651–666. https: //doi.org/10.1109/JPROC.2019.2895105

[9] Iftene, A., & Trandabă, D. (2020). Eye and voice control for an augmented reality cooking experience. Procedia Computer Science, 176, 1469–1478. https: //doi.org/10.1016/j.procs.2020.09.157

[10] Han-byul, J., Jang-Woon, K., Chil-Woo, L. (2007). Augmented reality cooking system using tabletop display interface. CEUR Workshop Proceedings, 260.

[11] Zhou, Q. W., Chen, H., & Gao, Y. (2022). Design and implementation of VR-based visual interactive system for Sichuan cuisine cooking. Modern Information Technology, 6(9), 19–21, 25.

[12] Bouck, E., Norwine, L., Jakubow, L., et al. (2025). Playing with food: Non-immersive virtual reality to teach food preparation. Journal of Developmental and Physical Disabilities, 1–19. https: //doi.org/10.1007/s10882-025-10028-6

[13] Junhao, X., JJC, K. J., Carol, O., et al. (2022). Trends of augmented reality for agri-food applications. Sensors, 22(21), 8333. https: //doi.org/10.3390/s22218333

[14] Reshma, K., Anusha, S. (2018). Real-time image tracking for augmented reality application. Journal of Computational and Theoretical Nanoscience, 15(6), 2471–2476. https: //doi.org/10.1166/jctn.2018.7491

[15] Janette, H., & Melanie, M. (2018). Developing 21st century competencies of marginalized students through the use of augmented reality (AR). LEARNing Landscapes, 11(1), 153–169. https: //doi.org/10.36510/learnland.v11i1.929