Design of a Pet Adoption Responsibility Assessment System Based on Spatial Computing and Environmental Semantic Understanding

Jiayue Ma

Royal Melbourne Institute of Technology University, Master of Design Innovation and Technology, VIC 3003, Melbourne, Australia
mjy020406@gmail.com

Abstract. With the rapid development of space computing technology, the Mixed Reality (MR) system shows great potential in understanding and enhancing the physical world. Aiming at the problem of the high abandonment rate caused by impulsive decision-making in current pet adoption, this paper proposes a pet raising simulation system based on MR technology. The system uses lidar scanning and AI scene recognition technology to integrate virtual life into users' real-life environment. By simulating the use of real space objects by pets and the stress response to environmental parameters, users can preview the life of pet care at home. This study aims to explore whether this embodied interaction can effectively evaluate and enhance the adopter's sense of responsibility, and transform the application of MR technology from entertainment activities to social behaviour intervention and evaluation tools. This study proposed the design direction of the scale of responsibility for raising pets as the core, and verified the technical feasibility of this path from the system design level, providing a new design reference for MR transformation from entertainment media to social behaviour intervention tools.

Keywords: Mixed reality, Spatial mesh, XR spatial interaction, Environmental perception, Pet adoption.

1 Introduction

In recent years, the concept of "adoption instead of purchase" has gradually become a social consensus, but the problem of "secondary abandonment" has become increasingly serious. Many adopters tend to have perfect psychological expectations when making decisions, ignoring the time cost and environmental damage caused by pets in real life, such as scratching furniture, excreting everywhere, and emergencies, such as illness and noise stress. The existing adoption review mechanism mostly relies on written questionnaires and oral commitments. This static review method is difficult to truly reflect the adopters' emotional management ability and sense of responsibility in the face of actual home conflicts. Therefore, a simulation assessment tool that can conduct situational stress testing in users' real-life spaces is needed.

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_67

At present, Mixed Reality technology has been widely explored in the fields of education, medical treatment and entertainment, and the relevant technical paths and development challenges are also reviewed [1]. The core of MR technology is to integrate virtual information with the real environment, so that users can interact with virtual objects more naturally, forming a new type of embodied interactive experience [2].

Unlike the on-screen virtual pet game, the MR head display has been able to generate a more accurate three-dimensional space grid with the help of lidar depth sensing and slam technology, support the occlusion and light and shadow synchronisation between virtual and real objects, and initially has the ability to understand environmental semantics. However, the existing systems still have obvious deficiencies in integrating AI semantic understanding with the behavioural logic of virtual life. It is difficult for virtual pets to make reasonable responses based on the functional attributes of real objects. Based on this, this study attempts to introduce the environmental semantic understanding into the MR pet raising simulation system as a bridge between the real home conflict and the psychological assessment of adopters.

In related research, scholars have discussed the application of Mixed Reality technology in behaviour intervention and situational simulation. Herrera's research shows that the immersive virtual environment has empirical support for the change of users' behaviour and attitude, and provides a reference for the cultivation of a sense of responsibility [3]. The 3D dynamic scene graph framework proposed by Rosinol et al. Realises the semantic modelling of the object function relationship, and lays the technical foundation for the semantic-driven behaviour of virtual objects [4]. Subsequent studies also further confirmed that the embodied experience of perspective conversion has a significant and sustained effect on improving empathy and prosocial behaviour [3]. However, there is still no research on the simultaneous integration of spatial semantic perception, virtual life behaviour logic and responsibility assessment in real residential space. This study is just an attempt in this direction.

This paper aims to fill the research gap between the current spatial scanning visualisation and social assessment, propose an interactive spatial environment perception and evaluation system, and apply it to the pet adoption scene. This study first explores whether the MR system combined with environmental semantic understanding can effectively reproduce the cross-species survival conflict in the user's real physical space, and then verifies whether this embodied preview experience can effectively change the user's idealised cognition of adoption, so as to be applied as a preemptive responsibility assessment tool.

2 Theoretical Foundations and Related Work

2.1 Spatial Computing and Environmental Semantic Understanding

At present, many research results on spatial reconstruction algorithms have been accumulated in the field of AR and MR, but most systems still remain at the level of geometric perception, and only treat real objects as simple physical obstacles. This study introduces the concept of functional visibility and further infers its functional attraction to virtual pets on the basis of identifying object categories. For example, the leather sofa is marked as a high-traction, attractive surface. This semantic mapping is the technical premise to drive the reasonable interaction between virtual pets and furniture.

2.2 Embodied Cognition and Empathy Construction

The Proteus effect points out that the impact of immersive virtual experience on an individual's attitude towards reality is quite significant [5]. In the field of animal protection education, perspective conversion is considered to be an effective means to help users understand the animal situation and establish empathy. This study introduces this theory into MR Design, and represents the daily home environment as an oppressive pet perception perspective by switching the user's perspective, so as to stimulate the adopters' more real empathy experience.

As far as embodied cognition and empathy research is concerned, the existing achievements focus more on the attitude change of public issues such as environmental protection and endangered species conservation, and the research on the application of embodied simulation technology in the pre-assessment of individual life responsibility is almost blank, especially the empirical exploration in the real living environment of users is extremely scarce. In view of this gap, this study puts forward two core assumptions: first, users who have completed the MR home care simulation experience will have a significant increase in the score of their sense of responsibility for caring; Second, the destructive events caused by the semantic interaction between virtual pets and real furniture will make users' score of potential abandonment risk significantly higher than the control group who did not experience such conflict situations.

3 System Architecture

In order to achieve the above goals, this study designed an MR evaluation system based on spatial computing and environmental semantic understanding. The system consists of three core modules: Environment Awareness, Behaviour Engine and Rendering Feedback. In terms of pet category setting, the system currently supports virtual behaviour simulation of cat and dog pets, which have significant differences in behaviour mode, space occupation and stress response. As far as cats are concerned, scratching furniture is an autonomous, natural behaviour. Research shows that more than 85% of the surveyed cat owners report that their pets have improper scratching problems with furniture, and cats are highly sensitive to environmental changes. Strange environments or sound stimuli often trigger strong stress reactions, which are manifested by hair erection, increased vocalisation, and avoidance behaviour [6]. As far as dogs are concerned, stress behaviour is the core indicator for evaluating the welfare status of dogs, including observable behaviours such as vocalisation, limb stiffness and abnormal interaction with the non-social environment. Its dependent interference with the host's work and rest rhythm is also one of the common home conflicts faced by adopters [7]. This kind of logic design of species' differentiated behaviour based on animal ethology enables the system to provide more realistic and targeted home conflict simulation for different adoption intentions. The overall architecture of the system is shown in Figure 1.

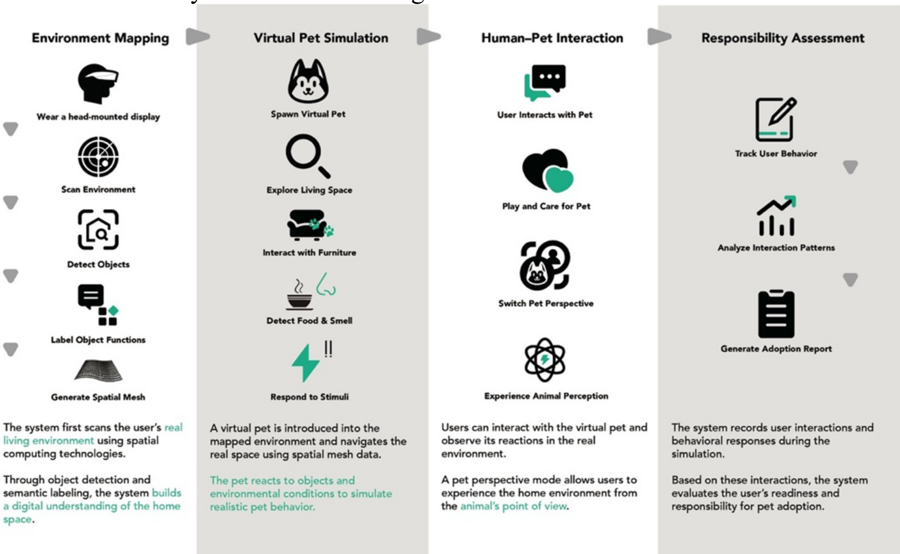

Fig. 1. Architecture of the MR Pet Adoption Responsibility Assessment System Based on Spatial Computing and Environmental Semantic Understanding. (Picture credit: Original)

The workflow of the system is divided into four stages: Environment Mapping, Virtual Pet Simulation, Human Pet Interaction, and Responsibility Assessment. The complete technical implementation path of the system is shown in Figure 2.

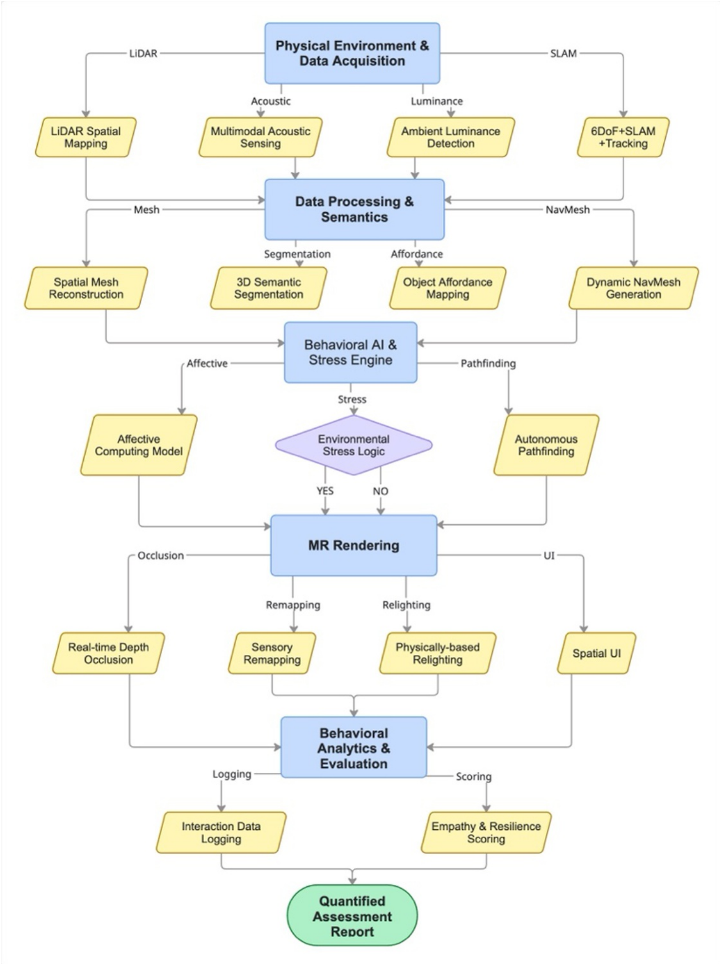

Fig. 2. Interaction Logic Flowchart of the MR Pet Adoption Responsibility Assessment System Based on Spatial Computing and Environmental Semantic Understanding. (Picture credit: Original.)

4 Key Technologies

4.1 Real-Time Physical Environment Perception and Meshing

In the MR system, reliable spatial modelling is the basis of virtual reality fusion interaction. Spatial computing systems usually rely on a visual SLAM algorithm to complete environment positioning and 3D modelling, so as to support the stable superposition of virtual objects in the real environment [8]. The real-time three-dimensional reconstruction technology can generate high-precision spatial grids, enabling virtual objects to physically collide with real objects and deep occlusion [9]. The above technologies ensure the autonomous path planning ability of virtual pets in complex home environments. Figure 3 shows the semantic annotation results of the system for daily home scenes, including the functional attribute identification of sofas, potted plants and air purifiers, and the detection of the interaction between the virtual golden retriever "Kiki" and family members. The status panel at the top right corner of the screen records the pet's trust value, energy value and pressure value in real time, accumulating behaviour data for subsequent responsibility assessment.

Fig. 3. Illustration of MR System Environmental Semantic Annotation and Real-Time Multi-Dimensional State Monitoring of Virtual Pets. (Picture credit: Original)

4.2 Spatial Semantics and Functional Affordance Mapping

In recent years, deep learning methods have been widely used in scene semantic understanding, which enables the system to recognise the categories of objects in space and infer their functional attributes. In order to enable virtual objects to make more complex behaviour decisions in the real environment, researchers proposed the concept of a dynamic scene graph, which realises semantic modelling of the environment by integrating spatial structure, object categories and functional relationships [4]. On this basis, the system introduces an interactive semantic annotation framework. Semantic segmentation technology can automatically identify different object categories in the environment and provide semantic information for environmental interaction [10]. In the initialisation phase, the system guides the user to use holographic rays to actively scan and mark the furniture and interactive status in the home. Thus, the pet behaviour conflict is driven by the semantic attributes of real objects, which significantly improves the ecological validity of the situation.

Figure 4 shows a typical interaction moment of the system in a real home scene: the system marks the dark fabric sofa in the living room as an object with high injury risk through spatial semantic recognition, and the virtual cat "yunyun" immediately triggers the scratching behaviour. The real-time stress panel at the upper right corner synchronously displays its status as "aggressive scratching", and the pressure value rises to a high level. At the same time, the system pops up a holographic evaluation interface, prompting the user that the current scene needs immediate intervention, and guiding him to make a yes/no decision through gestures, with intervention teaching content attached below. This design integrates semantic-driven conflict generation, real-time state feedback and evaluation nodes, which is the key to the system from visual overlay to behaviour evaluation tool.

Fig. 4. Schematic diagram of the interaction interface for assessing destructive behaviour and adoption of virtual pets in an MR system driven by spatial semantics. (Picture credit: Original)

4.3 Environmental Stress and Behavioural AI Engine

The system has built-in stress logic based on multimodal sensor data, and the environmental decibel value is collected in real time through the device's microphone array. In case of sudden high decibel noise, such as heavy objects falling to the ground and the sound of TV, the AI engine of the virtual pet will immediately trigger the anxiety state, which is manifested as hair erect, limb tremor, and will automatically move to the lower part of the physical cabinet in reality to hide. This mechanism directly converts the sound changes in the environment into the stress behaviour output of virtual pets.

Figure 5 shows the noise stress triggering effect of the system. When the environmental decibel reaches the set threshold, the virtual dog and the virtual cat enter the high-pressure stress state synchronously, the stress bar rises to the red warning range, and the pet perception label identifies children's toys and floor sweeping robots as noise threat sources. Then the system pushes the intervention prompt panel to users to guide adopters to take measures such as appeasement, diversion of attention or screening of stressors, so as to turn sudden noise events into a key node of responsibility assessment.

Fig. 5. Illustration of the MR System's Multimodal Noise-Induced Stress Triggering Mechanism and Pet Soothing Intervention Prompt Scenario. (Picture credit: Original)

4.4 Gesture Interaction System

In order to enhance the interactive experience of users in the Mixed Reality environment, the system introduces gesture recognition as the main input mode. Compared with traditional controllers, gesture interaction allows users to operate virtual objects in a more intuitive way, which helps to improve the immersion and ease of use of the system.

The system captures the user's hand posture in real time through the depth camera of the head-mounted MR device and the hand tracking algorithm, analyses the finger joint position and motion trajectory, and recognises basic gestures such as air tap, drag, grab, pointing, etc. Previous review studies have pointed out that the hand tracking natural gesture is the most mainstream input method in MR head display, and its direct sense of control helps to reduce the cognitive burden and improve the interaction naturalness [11]. Users can use these gestures to operate the virtual interface in the space, including interacting with pets, delineating behaviour areas, triggering interactive menus, etc.

4.5 Eye-Tracking Interaction System

The system obtains the user's line of sight vector in real time through the eye tracking module built in the MR device, and calculates the interactive object pointed by the line of sight, combined with the spatial grid model. When the user looks at a virtual interface or real object, the system will automatically highlight the target and give interactive prompts; If the user keeps looking at the virtual pet, it will trigger the pet's active response - turning towards the user or actively approaching. This vision-driven interaction mechanism provides a more natural experience dimension for the interaction between people and pets, which helps to enhance the sense of immersion.

4.6 Pet Perspective Mode

Immersive virtual experience can help improve users' empathy and affect their attitude towards social problems, which has been supported by research [12]. To help users more intuitively understand the perceived world of animals, the system adds the pet perspective mode, allowing users to switch to the perception perspective of virtual pets in the MR environment, and experience the same living space from the perspective of animals.

There are fundamental differences in perception between the pet perspective and the human perspective. Humans are used to visual dominance, and are often insensitive to the noise, smell and other stimuli around; And the same space in the perceptive world of animals may be full of potential threats, temptations and stressors. Switching to the pet perspective, users can jump out of their own experience framework and directly feel the different pressure experience of the same environment on different species. In this mode, the system visually reconstructs the perceptual information in the real environment: the food smell in the kitchen is presented in the form of floating particles or color flow, allowing users to intuitively feel that pets are highly sensitive to the smell; The noise generated by vacuum cleaners or electrical appliances is converted into vibration waveforms to simulate the stress state of animals in a high noise environment. The system can also restore the visual characteristics of different animals through colour filtering or contrast adjustment, such as the perceptual differences of cats and dogs on a specific colour gamut. Figure 6 shows the perceptual reconstruction effect after switching to the dog perspective. The screen tone has been filtered according to the dual colour visual characteristics of dogs.

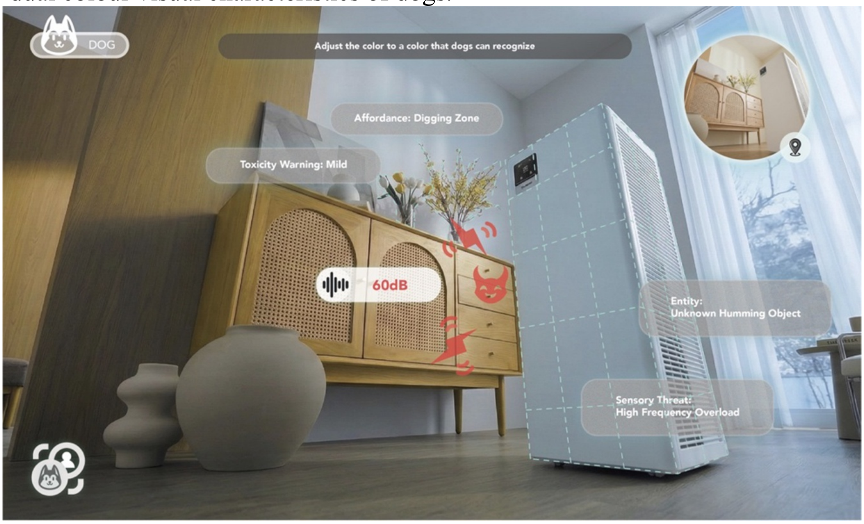

Fig. 6. Illustration of Environmental Perception Reconstruction in MR Pet Perspective Mode (Canine Perspective). (Picture credit: Original)

5 Conclusion

Focusing on the practical problem of second abandonment in pet adoption, this paper proposes and designs an MR Responsibility Evaluation System based on spatial computing and environmental semantic understanding. The system uses MR technology to introduce virtual pets with behavioural logic into users' real home scenes, breaking through the limitations of traditional written reviews. Through the elaboration of the interactive logic and technical path of spatial scanning, grid visual feedback, holographic UI design and object guidance annotation, this study shows that this design with visual semantic feedback and structured guidance can effectively improve users' awareness of the home space perception state. The two core research issues proposed in the introduction have been preliminarily verified at the technical level, but the verification effect still depends on the support of subsequent user experimental data, which is also the main task of the next stage of research.

Of course, this study still faces some practical challenges. The dependence of the system on hardware performance is obvious. The scanning accuracy of lidar and the effect of AI semantic recognition are limited by the computing power of the existing MR head-mounted display. In complex home environments, there may be object recognition errors or grid reconstruction errors. The behaviour logic of virtual pets is mainly driven by rules. When facing different living spaces, there is room for improvement in the sense of reality and scene adaptability of behaviour. In the assessment of sense of responsibility, the deviation of social expectations is also a potential interference factor. How to design a more objective behaviour data collection scheme is a key issue to be tackled in the future.

To solve the above problems, future research can be promoted in two directions: on the one hand, explore lighter end-to-side AI models to improve the real-time and accuracy of semantic recognition and reduce the dependence on high-end hardware; On the other hand, reinforcement learning mechanism is introduced to optimize the adaptive behavior logic of virtual pets to make them behave more naturally in more space scenes.

With the gradual lighting and popularisation of MR equipment, the system is expected to provide a set of digital assessment tools for pet adoption review, and its methodological framework also has the potential to extend to social behaviour intervention scenarios such as parenting preparation assessment and elderly care pre-assessment.

References

1. Rokhsaritalemi, S., Sadeghi-Niaraki, A., Choi, S.M.: A review on mixed reality: Current trends, challenges and prospects. Applied Sciences 10(2), 636 (2020).
2. Speicher, M., Hall, B.D., Nebeling, M.: What is mixed reality? In: Proceedings of the CHI Conference (2019).
3. Herrera, F., Bailenson, J., Weisz, E., Ogle, E., Zaki, J.: Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. PLoS One 13(10), e0204494 (2018).

4. Rosinol, A., Gupta, A., Abate, M., Shi, J., Carlone, L.: 3D dynamic scene graphs: Actionable spatial perception with places, objects, and humans. arXiv preprint (2020). https://doi.org/10.48550/arxiv.2002.06289
5. Yee, N., Bailenson, J.: The Proteus effect: The effect of transformed self-representation on behavior. Human Communication Research 33(3), 271–290 (2007).
6. Cisneros, A., Litwin, D., Niel, L., Stellato, A.C.: Unwanted scratching behavior in cats: Influence of management strategies and cat and owner characteristics. Animals 12(19), 2551 (2022).
7. de Winkel, T., van der Steen, S., Enders-Slegers, M.-J., Griffioen, R., Haverbeke, A., Groenewoud, D., Hediger, K.: Observational behaviors and emotions to assess welfare of dogs: A systematic review. Journal of Veterinary Behavior 72, 1–17 (2024).
8. Teed, Z., Deng, J.: DROID-SLAM: Deep visual SLAM for monocular, stereo, and RGB-D cameras. In: Advances in Neural Information Processing Systems 34, pp. 16558–16569 (2021).
9. Kähler, O., Prisacariu, V.A., Murray, D.W.: Real-time large-scale dense 3D reconstruction with loop closure. In: ECCV, pp. 1–17 (2016).
10. Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P., Garcia-Rodriguez, J.: A survey on deep learning techniques for image and video semantic segmentation. Applied Soft Computing 70, 41–65 (2018).
11. Nguyen, R., Gouin-Vallerand, C., Amiri, M.: Hand interaction designs in mixed and augmented reality head mounted display: A scoping review and classification. Frontiers in Virtual Reality 4 (2023).
12. Adhanom, I.B., MacNeilage, P., Folmer, E.: Eye tracking in virtual reality: A broad review of applications and challenges. Virtual Reality 27(2), 1481–1505 (2023).