**Design Process Report – Evaluation**

Peilin li s4797896

**Objective and validation metrics:**

In this prototype test, my primary goal was to verify whether the designed interactive features effectively support user operational needs in an immersive environment and to initially assess their advantages over traditional 2D interfaces. The specific objectives are as follows:

First, in terms of operational learning curve, this test aims to verify whether users can quickly understand and master interactions such as grabbing, zooming, and comparing. If users can accurately and independently complete the actions after brief instructions, with one or fewer errors, this objective is considered achieved.

Second, in terms of efficiency, the test will assess whether spatialized interaction improves the speed and fluency with which users complete tasks. If most users can complete tasks within the expected time (e.g., categorization tasks within 15 seconds), this indicates a positive impact on efficiency.

Third, in terms of accuracy and controllability, this test will focus on user performance when performing high-precision operations such as editing. If most users can accurately complete the actions at the exact location or time (selecting a comparison video allows users to accurately pause at a specific location) without requiring three or more attempts, this indicates a high degree of reliability. Fourth, and finally, regarding functionality and subjective experience, the test will collect users' subjective evaluations of immersion, intuitiveness, and ease of use through a questionnaire. An average score above 3 indicates that the design provides an acceptable or even good user experience.

**Results:**

Overall, all six participants were able to complete the tasks, but there were multiple errors or failures in the zooming and editing phase. The classification task generally took a short time, with some minor errors. The side-by-side comparison operation was generally completed with high completion rates, although some users experienced deviations during the pause phase. The questionnaire results showed that immersion scores were generally high, while intuitiveness and ease of use varied significantly.

**Analysis/Insights:**

A comprehensive analysis of the participants' test results revealed the following core themes:

Low learning cost and intuitive interaction

Partially achieved the objectives. After a brief explanation, most users were able to independently complete the grabbing, classification, and comparison operations, with a reasonable number of errors. However, due to insufficient prompts in the classification function, some users expressed doubts during use, indicating that guidance in this area still needs improvement. High-Precision Tasks Reveal Bottlenecks

In tasks requiring precise control, such as zooming and editing, users frequently made errors or failed, generally relying on multiple attempts. This reflects that while immersive spatial manipulation offers a sense of realism, it has inherent limitations in fine-grained control.

Experience Divergence

In the questionnaire, some users found the operation natural and smooth, while others were frustrated by insufficient precision or instability. This reveals that the concept has potential in terms of "immersion" and "intuitiveness," but still needs improvement in "reliability" and "controllability."

**Evaluation of Aims:**

Combined with the initial test objectives, the following evaluation conclusions can be drawn:

Operational Learning Cost: Target partially achieved. Most users were able to complete the operation independently, but the classification function lacked prompts, leaving some users confused.

Efficiency: Target partially achieved. Classification and zooming tasks were completed within the expected timeframe, but the comparison phase was inefficient.

Precision and Controllability: Target not achieved. Zooming and editing errors were frequent, making it difficult to meet precision requirements.

Subjective Experience: Target partially achieved. Immersion was rated highly, but intuitiveness and ease of use varied widely, and the overall results were less stable.

**Concept Iteration:**

Based on the aforementioned test results and target assessments, this iteration will focus on four key areas: functional optimization, efficiency, accuracy, and subjective experience. The following specific improvements are proposed:

Functional Optimization

To address the confusion some users still experience with grabbing and categorizing, the new iteration will introduce more intuitive interactive prompts. For example, when a user approaches a grabbable object, the system will automatically highlight the object and cover it with the mouse cursor, reducing initial uncertainty. Furthermore, the ability to categorize grabbable objects will be added, rather than simply categorizing different materials. Users can now categorize different projects based on their own ideas.

Efficiency Improvements

The compare function will introduce separate pause, rewind, and synchronized fast-forward functions, making it easier for users to compare specific locations and avoid wasted time and inefficiencies.

Improvements in Accuracy, Controllability, and Usability

To address errors that may occur during zoomed editing, the new iteration will add dynamic tick marks and a visual ruler, allowing users to clearly locate cut points on the zoomed timeline. Furthermore, feedback will be enhanced, with cutting effects and displays during the cut process to help users understand their actions.

**Reflection on the concept/design/methodologies/future testing and planning:**

This prototype testing provided a deeper understanding of the design process and methodology, and gained important insights from user feedback. Overall, the testing process helped me confirm the concept's potential value while also revealing many shortcomings in practical use.

At the conceptual level, the spatial editing approach proved effective in enhancing immersion and making asset management more intuitive, a point confirmed by multiple user feedback. However, when tasks demanded high precision, the limitations of the existing design became apparent. For example, the frequent errors in zoom editing demonstrated that relying solely on immersive spatial interaction was insufficient to support high precision requirements. This reminds me that future concept development should not solely pursue immersion; it must also strike a balance between precision and usability.

In terms of design methodology, the task-driven testing process helped me obtain more intuitive data, enabling me to clearly observe user behavior patterns and operational bottlenecks. However, the testing process also exposed methodological shortcomings: the relatively limited number of recorded dimensions lacked a deeper analysis of user failure causes, and the questionnaire relied too heavily on subjective ratings, failing to provide sufficient quantitative metrics. This suggests that future designs need to incorporate more diverse data collection methods, such as quantitative metrics like task success rate and margin of error, to more accurately measure design effectiveness.

In terms of learning outcomes, I've realized that prototype testing isn't just a functional test, but also an exploration of user cognition and operational habits. Users' efficient completion of classification and grasping tasks demonstrates that spatial interaction can reduce learning costs; however, repeated failures in high-precision tasks also made me realize that users rely more heavily on feedback mechanisms than expected.

In future testing and planning, I plan to expand the number and diversity of test participants in the next iteration, particularly including more users with video editing experience to obtain more professional feedback. I will also explore adding "Highlight Assist and Guidance" and "Keyframe Assist" features to the prototype, and incorporate more standardized comparative experiments into my testing methodology (such as a control group against traditional 2D interfaces) to more clearly evaluate the unique strengths and weaknesses of XR editing tools. In summary, this test not only allowed me to appreciate the potential of my design concept, but also made me aware of its shortcomings and room for improvement. Future design and testing will focus more on balancing immersiveness and precision in the user experience, and will utilize more systematic testing methods to validate and optimize the design.

**Appendices**

Test Records

In this test, six participants completed the task, with the following records:

Participant 1: Sorting was completed within 14 seconds with only one misplacement; zooming and cutting was inaccurate the first time, but successful the second time; side-by-side comparison required multiple pauses and selections. Questionnaire Ratings: Immersion 4/5, Intuitiveness 4/5, Ease of Use 3/5.

Participant 2: Initially unfamiliar with the operation, sorting took 18 seconds and resulted in two errors; zooming and cutting failed to complete accurately three times in a row; and side-by-side comparison required two attempts to confirm. Questionnaire Ratings: Immersion 3/5, Intuitiveness 2/5, Ease of Use 2/5.

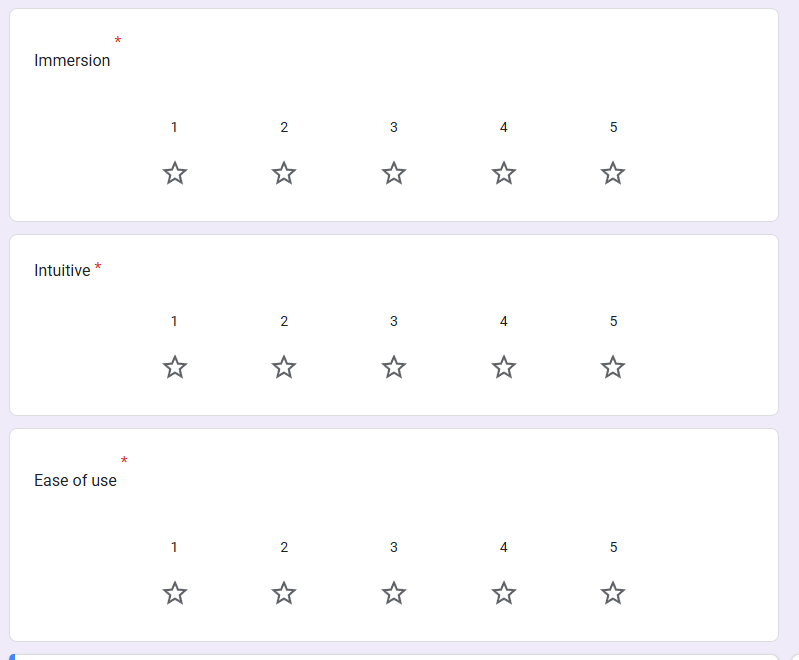
Participant 3: Operated independently with little to no instruction, sorting in 12 seconds; zooming and cutting completed accurately in one try; and side-by-side comparison completed pauses and selections smoothly. Questionnaire Ratings: Immersion 5/5, Intuitiveness 5/5, Ease of Use 4/5.

Participant 4: Sorting took 20 seconds, exceeding expectations; zooming and cutting completed smoothly; side-by-side comparison had a 1–2 second error. Questionnaire ratings: Immersion 3/5, Intuitiveness 2/5, Ease of use 2/5.

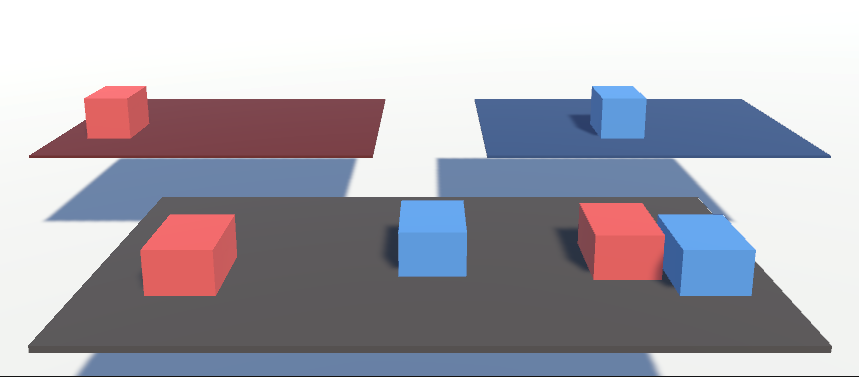
Participant 5: The categorization task took 15 seconds and was completed smoothly; zooming and clipping was completed smoothly; and the side-by-side comparison was smooth. Questionnaire ratings: Immersion 4/5, Intuitiveness 4/5, Ease of use 4/5.

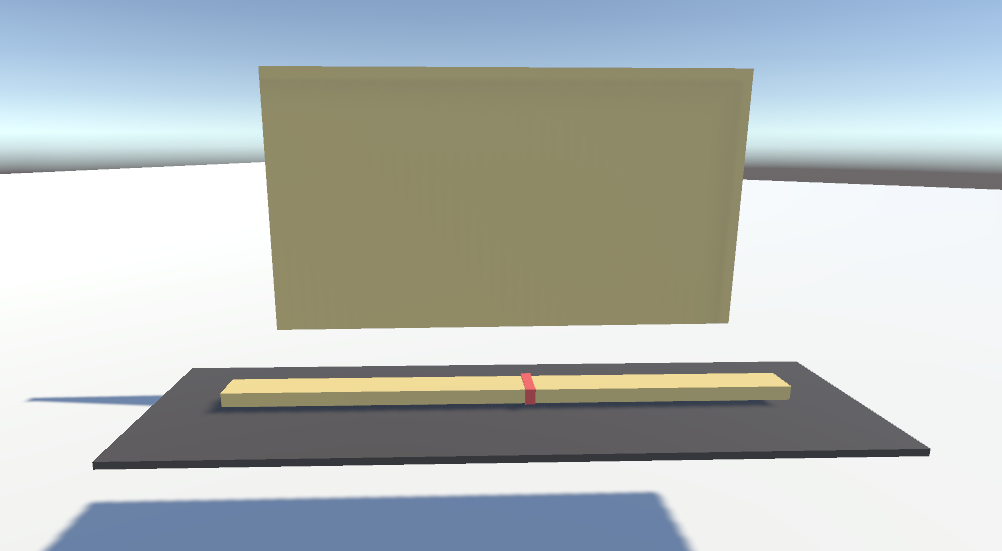
Participant 6: The categorization task took 13 seconds and had no obvious difficulty; zooming and clipping was difficult to find the cut mark, but was successful the second time after being reminded; and the side-by-side comparison was able to accurately pause and select on the second try. Questionnaire ratings: Immersion 5/5, Intuitiveness 4/5, Ease of use 3/5.

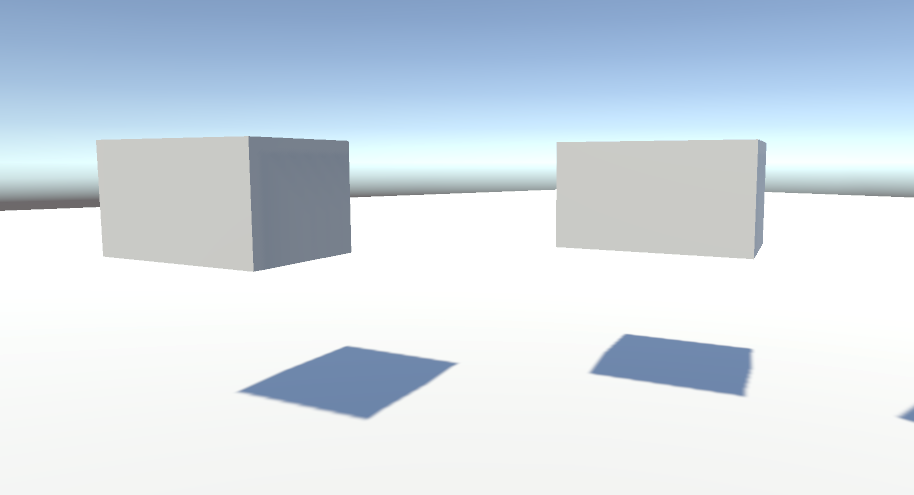
Rating Sheet



Prototype screenshots







Statement of Originality

I confirm that the Unity prototype design submitted here is my original work. During development, I used AI tools to assist in coding some interactive features, and I used Google Translate for translation and grammar checking. Except as noted above, all Unity implementation and textual content are my own.