Coordinated 3D Interaction in Tablet- and HMD-Based Hybrid Virtual Environments

Jia Wang
HIVE Lab
Worcester Polytechnic Institute
wangjia@wpi.edu

Robert Lindeman
HIVE Lab
Worcester Polytechnic Institute
gogo@wpi.edu

ABSTRACT

Traditional 3D User Interfaces (3DUI) in immersive virtual reality can be inefficient in tasks that involve diversities in scale, perspective, reference frame, and dimension. This paper proposes a solution to this problem using a coordinated, tablet- and HMD-based, hybrid virtual environment system. Wearing a non-occlusive HMD, the user is able to view and interact with a tablet mounted on the non-dominant forearm, which provides a multi-touch interaction surface, as well as an exocentric God view of the virtual world. To reduce transition gaps across 3D interaction tasks and interfaces, four coordination mechanisms are proposed, two of which were implemented, and one was evaluated in a user study featuring complex level-editing tasks. Based on subjective ratings, task performance, interview feedback, and video analysis, we found that having multiple Interaction Contexts (ICs) with complementary benefits can lead to good performance and user experience, despite the complexity of learning and using the hybrid system. The results also suggest keeping 3DUI tasks synchronized across the ICs, as this can help users understand their relationships, smoothen within- and between-task IC transitions, and inspire more creative use of different interfaces.

Author Keywords

Hybrid virtual environments; 3D user interface; Tablet interface; Transitional continuity; Virtual reality

ACM Classification Keywords

H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems – *artificial, augmented, and virtual realities*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *evaluation/methodology, input devices and strategies, interaction styles, user-centered design*.

INTRODUCTION

Immersive virtual reality (VR) technology has been gaining great popularity recently thanks to a new generation of low-cost Head-Mounted Displays (HMD). Besides the high fidelity of the displays, the performance and usability of 3D User Interfaces (3DUIs) also play a critical role in the overall immersive experience delivered to the end user. Through decades of research, various input devices and interaction techniques have been proposed and evaluated for the basic 3DUI tasks of navigation, selection, manipulation, system control, and symbolic input [5]. But despite the realistic experience of grabbing and manipulating a virtual object using your hand [23], or real walking in a Virtual Environment (VE) [34], researchers also realize that interaction in VR can be just as confusing, limiting, and ambiguous as in the real world, when it comes to tasks with diverse requirements [28]. For example, it is difficult to select and manipulate objects of different sizes, from multiple angles, and at different distances, without spending significant time and effort on navigation.

One way to overcome such limitations is to develop Hybrid Virtual Environment (HVE) systems, which incorporate multiple and complementary virtual and/or physical interface elements appropriate for a set of tasks. For example, the World-In-Miniature (WIM) interaction technique renders an interactive miniature world in the left hand of the user to complement the immersive context with quick teleportation, range-less object selection, and large scale object translation [28]. HVE systems with different physical interfaces are inspired by Hybrid User Interface (HUI) systems [13]. A common example is the pen-and-tablet interface which uses a tracked surface to complement the spatial pen input for 2D tasks such as system control, symbolic input, and map-based way-finding [6].

The rapid progress of mobile technology has inspired a recent research trend of offloading 3DUI tasks to mobile phone and tablet devices, to take advantage of their growing computing power, high resolution, multi-touch touch screens, and various built-in motion sensors [4, 26, 33]. However, most of these techniques have been focused on very simple scenarios, where only one or two UI functions are assigned to the tablet to aid the primary spatial interface used in the immersive environment. Few studies have been conducted to investigate the overhead involved in transitioning between the multiple interface elements [14].

SUI'14, October 4–5, 2014, Honolulu, HI, USA.

http://dx.doi.org/10.1145/2659766.2659777

In this paper, we propose a novel HVE system that aims to join the strengths of a tablet device and an HMD-and-wand-based immersive setup. Instead of a supplementary tool, the tablet is designed and implemented as a complete Interaction Context (IC), formally defined later, which renders the entire virtual world on its own, and supports all 3DUI tasks through multi-touch gestures and 2D GUI elements. To reduce the perceptual, cognitive, and functional overhead [12] caused by complex 3DUI transitions across multiple ICs, a *coordination mechanism* featuring 3DUI task synchronization is proposed. Lastly, the results of a user study are presented, which suggest that task synchronization can lead to smoother transitions across ICs, and that user performance can be increased by using multiple complementary ICs in an HVE system.

RELATED WORK

Tablet-Based 3D Interfaces

Interactive tablets have been demonstrated as powerful tools for interaction in VR. By displaying an interactive 2D map on a tracked touchpad, early pen-and-tablet prototypes made way-finding and travel efficient in cluttered indoor spaces [1], as well as in large-scale outdoor scenes [6]. The Personal-Interaction-Panel (PIP) proposed concepts of a hybrid approach for object selection and manipulation, system control, and interaction with volumetric data [29]. The main idea was to augment virtual objects with 3D widgets and 2D GUI elements on the tablet, both of which could be interacted with using a stylus. Transparent pen and pad props have also been developed to enable Through-The-Lens (TTL) interaction with virtual content displayed on a tabletop [24]. From a usability point of view, an empirical study of a UI manipulation task has shown that bimanual interaction and passive haptic feedback offered by a physical surface held in the non-dominant hand can significantly increase precision and efficiency, as well as reduce fatigue [16]. Based on these advantages, the design guideline of *dimensional congruence* was proposed, which advocates matching the dimensionality of the 3DUI tasks to that of the input devices [11].

With no tethers attached, mobile phone and tablet devices can provide more flexibility than traditional pen-and-tablet interfaces. The use of mobile devices in VR has grown with the advancement of mobile technologies. Early work of Watsen *et al.* demonstrated a handheld computer used as an interaction device, which only contained simple 2D GUI widgets to aid system control tasks in the VE [32]. As the computing power increased, researchers started to experiment with rendering interactive virtual objects on the screen of mobile devices, based on PIP [4] or TTL [17] metaphors. Recently, many mobile devices contain high-performance, multi-touch touchscreens. To take advantage of this, various 3D interfaces have been proposed that combine multi-touch gestures with spatial tracking of mobile phones or tablets for object manipulation [33], volume data annotation, and textual data visualization [26]. Furthering this trend, a different design perspective is taken in this paper, which treats the mobile device not as a supplementary tool, but a complete interaction system, with computing power, display technology, and interaction richness comparable to that of an HMD-based, immersive VR system. This new approach is also expected to inspire new design possibilities of HVE systems for handling complex and highly diverse interaction tasks more effectively in 3D spaces.

Hybrid Virtual Environments

The early seminal work of Feiner & Shamash defined the term HUI as interface systems that combine heterogeneous display and interaction devices in a complementary way to compensate for the limitations of the individual devices [13]. Like HUI, HVE systems also strive to seamlessly integrate multiple representations of the same VE, in order to facilitate 3D interactions from different angles, scales, distances, reference frames, and dimensions. The multiple VE representations in HVE systems are often related based on some natural metaphor. For example, the WIM technique combines an egocentric and an exocentric view of the virtual world through a “handheld miniature world” metaphor [28]. The Voodoo Dolls technique creates a second instance of a remote object in the local space following a well-known fictional metaphor [20]. The SEAMs technique defines a portal which can be traveled through, or reached in to, to translate objects across two distinct spaces [25]. The Magic Lenses adopts an x-ray see-through metaphor to offer different visualizations of the same virtual content side by side [30].

HVE systems can also incorporate different physical interface components alongside the VE representations. The HVE system presented in this paper coordinates two VE representations contained in two ICs: a tablet device with multi-touch input and a 2D GUI, and an HMD-based VR system with wand input. Two closely related works are the HybridDesk, which surrounds a traditional desktop computer with a desktop CAVE display [9], and SCAPE, which puts a see-through workbench display in the center of a room with projection walls [7]. However, the former limited its ICs to exclusive 3DUI tasks, forcing the user to make unnecessary switches, and the latter mainly focused on view management, instead of rich 3D interactions.

Much research work in transitional user interfaces and Collaborative Virtual Environments (CVE) is closely related to HVEs. Transitional user interface systems present multiple representations of the virtual world in a linear, time-multiplexed way [14]. The MagicBook is a classic demonstration of a transitional experience between an exocentric view of the VE in Augmented Reality (AR) to an egocentric view represented in immersive VR [3]. Many CVEs can be considered as HVEs with their multiple VEs assigned to different users. A well-known metaphor is the combination of a God-user and a Hero-user, who possess complementary views and reference frames in the shared VE to aid each other towards a common goal [15]. The unique challenge of designing CVE systems is to ensure the collaborators are well aware of each other's viewpoints and interaction intentions as tasks are carried out, and avatars and artificial cues have been found effective [10]. Finally, it is also possible to merge hybrid, transitional, and collaborative virtual environments together into a hybrid collaborative system, such as the VITA system [2].

Cross-Context Transitions

Compared to traditional VR, one main challenge for HVE systems is the perceptual, cognitive, and functional overhead induced by transitions across multiple virtual and physical components [12]. The challenge is also present in coordinated multiple view (CMV) systems, where multiple views of the same dataset are generated and displayed to help the data analyst discover unforeseen patterns. The key to reduce the transition gap in CMV systems is to coordinate the visualizations of, and the interactions with, the multiple views [31]. For example, multiple views can be "snapped together" to better reveal their relationships and ease the gap between transitions [19]. Multiple views of 3D data can also be linked [22], or integrated through frame-of-reference interaction [21]. Guidelines for view management have been provided to minimize the cognitive overhead of context switching [31]. Applications and study results have demonstrated improvements in user performance when coordination mechanisms are implemented [27]. These findings inspired us to design and develop coordination mechanisms that can keep the complex 3D interaction transitions simple and smooth in the proposed HVE system.

METHODOLOGY

HVE Level Editor

Level editing was selected as the test bed to drive the design and study of our HVE system. It was selected for several reasons. First, level editing plays a key role in many real world applications, such as video game design, animation production, and urban planning. Second, many level-editing tasks feature diverse and complementary requirements, which makes them good candidates to adopt HVE approaches [6, 27]. Third, unlike the simple and monotonous tasks most VR studies have been designed for (e.g., travel from A to B [34]), level editing actually involves all 3DUI tasks (i.e., navigation, selection, manipulation, system control, and symbolic input) and combines them in various ways. This grants us an opportunity to study complex *3D interaction transitions across multiple ICs, and the overhead involved in the process*. The specific level-editing tasks supported in the proposed HVE system include editing of terrain (height and texture), foliage (grass and trees), objects, time-of-day, and spotlights.

Interaction Context

We introduce the concept of an Interaction Context (IC) here to represent *a conceptual integration of input and output devices, techniques, and parameters, which offers one representation of the VE and a set of interaction rules*. HVE systems are formed by relating multiple ICs under a unified metaphor. The metaphor defines the conceptual relationship between the ICs, making it more likely for the user to consider the overall HVE system as an integrated whole. Common HVE metaphors include WIM [28], portal [25], Voodoo Doll [20], see-through [30], and information surround [13]. For our HVE level editor, we selected WIM as the metaphor to combine the exocentric God view with the egocentric first person Hero view. An IC can be formed by specifying the following components:

- **Medium:** The type of medium adopted by the IC on the reality-virtuality continuum [18], such as VR, AR, or mixed reality.
- **Display device:** The multi-sensorial devices used to display the virtual world to the user's sensory organs, such as HMD, CAVE, headphones, haptic stylus, etc.
- **Rendering technique:** The technique used to represent the virtual content (e.g., shaders for visual display).
- **Input device:** The device used to express commands, such as a data glove or a multi-touch touch pad.
- **Interaction technique:** The software that maps the input data to control parameters in the virtual world. For example, wand input devices usually uses ray-casting based interaction techniques [23].
- **Perspective:** The position, orientation, and other parameters of a virtual camera that determines the IC's view of the virtual world. Immersive VR systems usually offer an in-the-world, first person perspective.
- **Reference frame:** The coordinate system that determines the perception of the virtual world and the effect of interaction. Egocentric (body-centered) and exocentric (object-centered) are two reference frames commonly discussed in VR [21].

This list of components defines a taxonomy that can be used to categorize HVE systems. For example, the original WIM interaction technique includes two ICs [28]. Both ICs use VR as the medium, and render their views of the VE in the same HMD, using a photorealistic shader. In addition, a buttonball prop is used in both ICs to interact with virtual objects, using a collision-based pick-and-drop technique. However, the two ICs are different in their perspectives and reference frames. The immersive IC has an in-the-world, first person view where all interactions are based on the user's egocentric body, while the miniature IC adopts an above-the-world, God view with object-centered exocentric reference frame. The HVE level editor presented in this paper incorporates an immersive IC and a tablet IC, whose components are specified in Table 1.

Components	Immersive IC	Tablet IC
Medium	Virtual reality	Virtual reality
Display device	HMD, fans	Tablet screen
Rendering technique	Photorealistic	Photorealistic
Input device	6-DOF wand	Touch screen
Interaction technique	Ray-casting & button based	2D GUI and multi-touch gestures
Perspective	In the world	Above the world
Reference frame	Egocentric (body-centered)	Exocentric (object-centered)

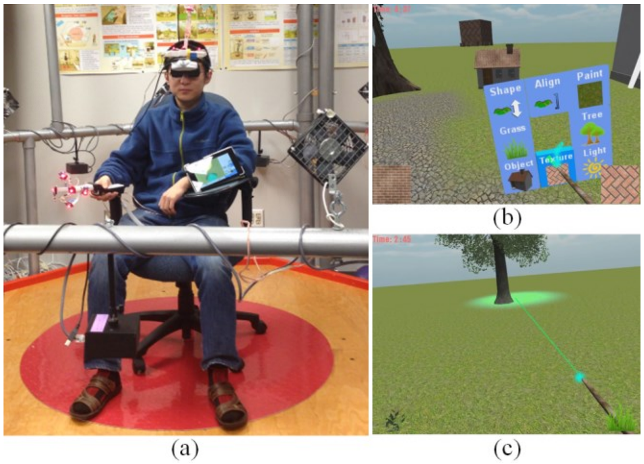
Table 1. The IC components of the HVE level editor

Immersive IC

As shown in Figure 1, an eMagin Z800 HMD is used to display a first-person, in-the-world view of a photorealistic VE, with a 60-degree horizontal field-of-view (FOV). The HMD utilizes two 800x600 OLED screens to render monoscopic images to both eyes with a 40-degree diagonal FOV. It is tracked in six degrees of freedom (DOF) using the PhaseSpace motion capture system. A constellation of four active LED markers is attached to the top of the HMD and tracked by sixteen cameras surrounding an octagon-shaped cage space, with the user seated in a swivel chair in the center. Since the HMD is non-occlusive, the user is able to see the display in the center of his/her field of view, as well as look at the screen of the tablet by gazing down.

A wand interface is provided to the dominant hand of the user to enable 3D interaction in the immersive VE. The wand is made by attaching a 6-DOF tracking constellation to a Wii Remote controller. 3DUI tasks are performed by pointing the wand and pressing buttons to issue commands. To navigate within the VE, the user can point the wand in different directions, and press down the D-pad buttons to travel in that direction at a constant speed. To reserve the realistic feeling, virtual locomotion is always constrained to the ground, but the swivel chair gives extra flexibility to point the wand easily at all directions. While the user is traveling, a group of fans corresponding to the direction of the locomotion are turned on, and blow wind at a constant speed to enhance the sense of motion in the virtual world.

To select an editing mode, the user can call out a floating menu as shown in Figure 1b, by holding down the “home” button on the Wii Remote controller. The tile pointed to by the wand is highlighted, and the corresponding editing mode is selected upon release of the “home” button. In the modes of terrain shape, texture, grass, or tree editing, a ray is cast from the tip of the wand to the intersection on the terrain surface, and a terrain brush is visualized to indicate the effective range. The size of the terrain brush can be changed using the “+” and “-” buttons on the wand controller. The “A” and “B” buttons have opposite effects. The former is used to raise, align, and plant trees and grass, while the latter is used to lower, sample, and remove trees and grass. In object editing mode, the objects in the VE, such as houses, can be selected by ray-casting and pressing the “A” button, or deselected by pressing the “B” button. Objects are highlighted in light blue when being pointed at, and in bright blue when actually selected. Once selected, the user can drag the object on the terrain surface by holding the “A” button, rotate it around the up-axis by pressing the left and right buttons on the D-pad, or scale it by pressing the “+” and “-” buttons. Lastly, the user can paint subparts of the virtual objects with different textures, as well as changing the scale of each texture.



(a) (b) (c)

Figure 1. The hardware setup (a), the floating menu (b) and terrain brush (c) of the HVE level editor.

Tablet IC

Figure 1a shows a user wearing a Google Nexus-7 tablet on his left forearm, and resting it on an arm pad to reduce fatigue. To leverage bimanual interaction [16], the user is asked to hold the wand interface temporarily in the left hand, or place it between the legs, and use the right hand to apply multi-touch gestures to the touch screen.

The interface on the tablet is illustrated in Figure 2. It consists of a three-tier GUI menu, a WIM view of the VE, and a shortcut bar. The top tier (1) is a tool bar for switching between the general editing modes. The tool bar at the second tier (2) displays further sub-modes, such as height, texture, grass, and trees for terrain editing. Based on the selection in the first two tiers, the third tier (3) shows specific GUI elements that can be used to perform the current task, such as a slider to resize the terrain brush, a selection grid to choose a type of grass to plant, and a broom button to clean grass from the terrain. Note that the immersive IC and the tablet IC each have their own terrain brush, so that terrain editing can be performed at different scales. To the right of the third-tier panel, an above-the-world, photorealistic, third person view of the VE is presented (4), whose camera has a 60-degree horizontal FOV in the VE, and can be manipulated using multi-touch gestures. These include a pinch gesture for zoom, a rotate gesture for orbit, a two-finger all-direction swipe gesture for pan, and a three-finger up-and-down swipe gesture for pitch. The one finger tap and swipe gestures are reserved for level editing, such as painting the terrain, or dragging an object on the terrain surface. The functionality of the shortcut buttons (5) will be discussed later.

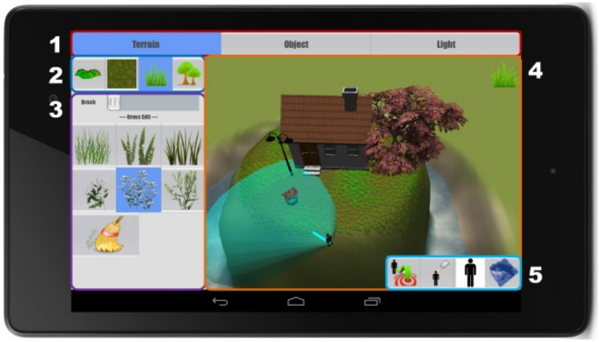


Figure 2. The tablet IC used to edit the VE from the God view

Regarding the software implementation, the HVE system was developed using the Unity game engine as a multi-player game running separately on the desktop and the tablet platforms. The hardware devices of the immersive IC are connected to the desktop computer through USB and Bluetooth connections. The input data from the PhaseSpace motion capture system and the Wii Remote controller are collected and streamed to the game process through VRPN and the Unity Indie VRPN Adapter (UIVA). Both the desktop and the tablet simulate the VE locally, and keep each other synchronized by sending UDP data streams and RPC calls over a local WiFi network. This way, both ICs can run the game at a steady 30 frames per second, and editing performed in one IC can be propagated to the other IC in real time, giving the user a convincing experience that they are viewing and interacting with the same virtual world, only from two different perspectives.

Coordination Mechanisms

The advantages of the two ICs can complement each other to support diverse tasks efficiently. For example, a fast way of moving a small object across a long distance in the VE is to select the object in the local space using the wand, and drag it to the destination using the tablet. However, such process involves frequent switches between the ICs, and the mental overhead of adapting to different IC components cannot be overlooked. The challenges to create smooth transition experiences in the HVE level editor are further illustrated in Figure 3, in which each level-editing task is decomposed into a set of basic 3DUI tasks. The user's workflow may start with any task in one IC and end with another task in a different IC. During transitions, the user needs to understand the relationship between the two VE representations, and adapt to distinctly different display devices, input devices, interaction techniques, reference frames, and perspectives. To reduce this transition gap, we propose the following four coordination mechanisms.

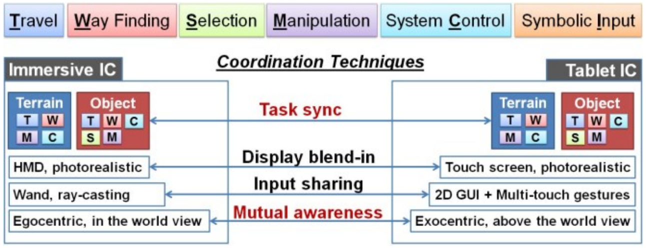


Figure 3. The coordination mechanism to smooth the complex cross-task, cross-IC transitions in the HVE level editor

- **Task synchronization:** The multiple data views in CMV systems are often coordinated to be consistent during user interaction [19, 22, 31]. Similarly, the effect of 3D interaction in one IC should also be propagated to all other ICs, to keep the workflow continuous during transitions. For example, when a user changes to object editing mode and selects an object using the wand, the tablet should also update to the same mode and select the same object, so that the user can directly continue to manipulate this object after changing the IC. Without task synchronization, the user's work would be interrupted, forcing her to repeat actions already made in the other IC.

- **Display blend-in:** The change of display device can cause perceptual gaps between ICs due to differences in screen size, resolution, brightness, and other parameters. Using mixed reality technology [8], the image of one IC's display device can be embedded into another IC's view to reduce this discrepancy. For example, compared to viewing the tablet screen from the peripheral vision, a better experience may be promised by tracking and rendering a virtual tablet in the HMD view, in place of the physical tablet itself.

- **Input sharing:** Some generic input devices, such as the mouse and keyboard, can be optimal to use in multiple ICs [2]. For example, a similar HVE system can be formed using a desktop computer and a tablet. In this situation, the mouse and keyboard could be efficient tools for controlling both the first-person view on the monitor and the God view on the tablet. Sharing input among ICs may not only reduce the mental overhead of transitions between interfaces, but also the physical effort of switching between devices.

- **Mutual awareness:** Research in CVE systems has stressed mutual awareness as the key to efficient human collaborations in VR [10, 15]. This rule can also be applied to HVE systems where different views are assigned to the same user. By knowing the whereabouts of the other view and the status of its interfaces, the user can better determine when to make the IC transition, and be more prepared to adapt to the new IC once the transition is made. Examples of effective mutual awareness cues include avatars, viewing frusta, pointing rays, and editing brushes (see Figure 4).

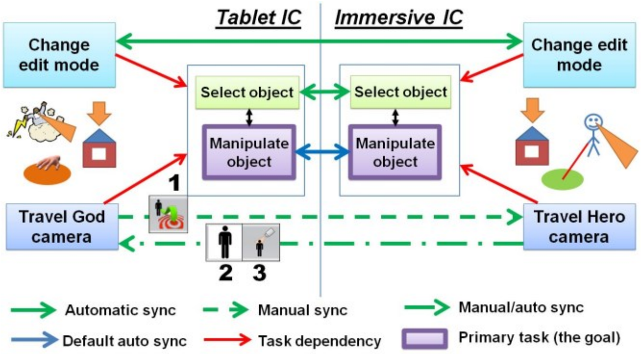


Figure 4. An example of task synchronization and mutual awareness cues implemented in the HVE level editor

Of the four coordination mechanisms, task synchronization and mutual awareness cues have been implemented in the current version of the HVE level editor. Figure 4 shows an example of the implementation in object-editing mode. The ultimate goal of this mode is to properly arrange virtual objects in the scene, through manipulation of the objects' positions, orientations, and scales. Manipulation is preceded by enabling object-editing mode (system control), moving to an appropriate spot (travel), and selecting the objet (selection). By default, the effect of object manipulation is synchronized between the two ICs, as the VE needs to look the same on both displays. However, synchronization of the preceding steps is optional, and very much dependent on the level of multi-tasking a hybrid system aims to support. We hypothesize that by synchronizing the effects of all 3DUI basic tasks, the working-memory demands required to keep track of the status of 3D interactions across ICs can be effectively reduced, leading to better task performance and user experience. Thus, task synchronization was implemented, with the goal of minimizing the interaction gap between the ICs. As illustrated in Figure 4, changing the editing mode or selecting a virtual object in one IC is always automatically synchronized to the other IC. Teleporting the user's Hero avatar to the field of the God view is done manually with the tap of a shortcut button (1) on the tablet, because previous research has indicated that constantly changing an immersive view can cause disorientation and even motion sickness symptoms [28]. To synchronize the God view with the space surrounding the Hero avatar, the user can either tap a button (2) for one-time teleporting, or switch a toggle (3) to enable/disable camera following.

EVALUATION

Hypotheses

The HVE system aims to combine the strengths of an immersive VR setup and a multi-touch tablet device. Being inside the virtual world, the user can better understand the space, judge scales of objects, and do manipulation of finer details [15]. Meanwhile, from the God view, the user can better navigate the VE, investigate the overall layout, and perform large-scale manipulations [28]. The two ICs are unified under the WIM metaphor, and coordinated through mutual awareness cues and task synchronization. Based on these analyses, we made the following hypotheses. ***H2*** and ***H3*** are trying to capture higher-level processes, such as user behavior, as opposed to low-level, performance-based claims as in ***H1***.

H1: Having the effects of basic 3DUI tasks synchronized between the ICs can make the transitions more continuous, and lead to better task performance and user experience.

H2: The users are able to learn the HVE system, and use both ICs to handle tasks with diverse requirements.

H3: The users are able to decompose a complex, high-level task into a series of basic 3DUI tasks, and find step-by-step strategies to efficiently use both ICs.

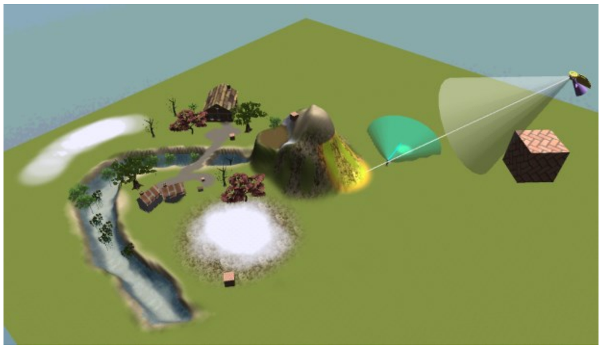


Figure 5. The task is to fix design flaws in an unfinished VE.

User Study

Instead of building a virtual world from scratch, the study presented the subjects an unfinished virtual world (see Figure 5), and asked them to find and fix five different types of design flaws in the VE as quickly and precisely as possible. This task approach was chosen for several reasons. First of all, based on natural metaphors, the design flaws were clear to identify, and the goals easy to understand and remember. Secondly, compared to building a VE from scratch, fixing existing design flaws takes less time to complete, making the threats such as user fatigue and motion sickness much more manageable. Finally, to complete the tasks efficiently, the subject needed to take different angles, interact at different scales and reference frames, and use different interfaces. This encouraged the subjects to learn both ICs, and explore different ways to use their complementary advantages.

With approval from the institutional review board (IRB), 24 university students were recruited with no remuneration. The study employed a within-subjects approach to compare the HVE level editor with and without task synchronization (indicated by green lines in Figure 4). The study began with the subject reading and signing the consent form, followed by a demographic questionnaire that asked about gender, age, and handedness, as well as experiences with immersive VR, multi-touch devices, multi-screen devices (e.g. the Nintendo WiiU), and first-person world building games (e.g., Minecraft). The subject was then introduced to the hardware used in the study, including the HMD, the wand, the tablet, and the fans. While having the freedom to swivel the chair, the subject was asked to stay in the center of the cage, to keep the best tracking quality of the motion capture cameras. The experimenter also explained the five world-fixing tasks as illustrated in Figure 6. The subject then put on the equipment, and learned the interfaces and the tasks in a 20-minute training session. To guide the subjects effectively, the VE in the training session had the five types of design flaws and the goals shown side by side as in Figure 6, where the experimenter explained different ways of solving each task, using either the wand or the tablet.

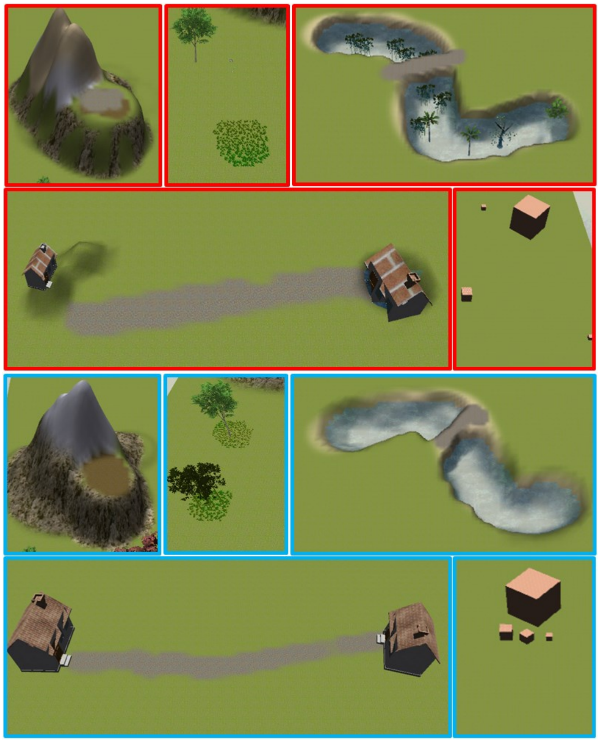


Figure 6. The five types of design flaws to fix in the study.

After the training session, the subject took a five-minute break, and then continued through two experimental conditions, each of which had one trial of world editing tasks. The conditions were presented to the subject in counterbalanced order, and only one of them had task synchronization enabled. To get used to the HVE system with different configurations, the subject spent eight minutes in a practice scene prior to each trial. In each trial, the subject had up to 15 minutes to fix the virtual world, and could end the trial early when they felt all design flaws had been addressed. After completing both conditions, the subject was asked to fill in a questionnaire to compare the HVE level editor with and without task synchronizations enabled, and to rate them on a one to six scale regarding eight different questions (see Figure 8). In the end, the subject was interviewed to give comments about the benefits and drawbacks of having multiple ICs, and the effectiveness of task synchronization.

Results

Task Performance

At the end of each trial, the system recorded the total time spent, and saved the edited VE into a data file. All VE data files were then reloaded and rated by two graders, who followed the same rubric to compare the completed VEs with the goals. The inter-rater reliability was evaluated using Pearson's correlation analysis and the result showed high agreement (R=0.92). As indicators of task performance, the task time, task score, and score-per-minute of the two conditions were compared using two-sided, paired t-test, with a threshold of 0.05 for significance. Score-per-minute was calculated by dividing score by time, and used as a measure of user efficiency. As indicated in Figure 7, subjects spent less time, and achieved higher task completeness, with task synchronization. The results are statistically significant for score-per-minute (p=0.02), and showed trends for task time (p=0.08) and score (p=0.07).

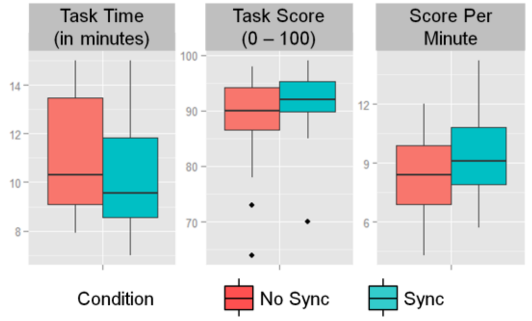


Figure 7. The analysis results of task performance indicators

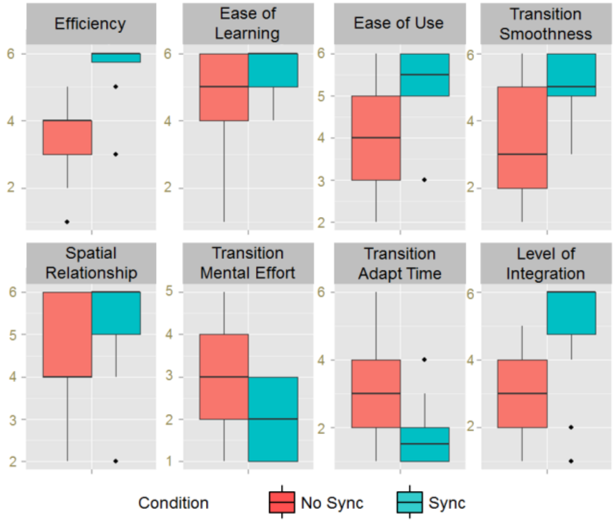


Figure 8. The analysis results of subjective rating scores

Post Questionnaire

The six-point rating scores of the two conditions were analyzed using two-sided Wilcoxon signed-rank tests with a threshold of 0.05 for significance on all questions. As indicated in Figure 8, the HVE system with task synchronization was considered to be more efficient, easier to learn, and easier to use, and the transitions between ICs smoother, and less time and mental effort demanding. In addition, the subjects felt the task synchronization mechanisms made it easier to understand the spatial relationship between the two VE representations, and the ICs were better integrated in the HVE system. All results were strongly statistically significant ($p < 0.01$).

Interview Feedback

In the interview, subjects were asked about whether they felt perceptual, cognitive, or functional disconnections between the ICs when transitions were made. The summary of their answers indicated better transitional continuity when task synchronization was enabled. The number of subjects who reported disconnected experiences, comparing “Sync” with “No-Sync”, were 6 and 11 for perceptual disconnection, 1 and 7 for cognitive disconnection, and 2 and 16 for functional disconnection. For the “Sync” condition, eight subjects complimented the synchronization of the editing mode, for emphasizing strong connection between the ICs, and making sure the non-active IC always kept up with the user’s workflow in the active IC. The travel synchronization buttons on the tablet (teleport, focus, and follow) also had significant contributions to the smooth transition experiences, according to eight subjects who claimed that “the two views were spatially connected with these buttons” and that “the appropriate camera view was always available at hand when I tapped these buttons”. Synchronization of selected objects was also liked by four subjects, as it enabled effortless within-task transitions, such as picking up a small cube using the wand and dragging it across the virtual world on the tablet screen. For the “No-Sync” condition, seven subjects felt the ICs were disconnected, and the overall HVE system was confusing and awkward to learn and use. Because the editing mode and the selected object did not get updated in both ICs, the subjects had to keep track of their individual status, and repeat actions they already took before the transitions. Four subjects even gave up using both ICs, and stayed with one interface throughout the trial. However, four subjects did point out one advantage of working in the “No-Sync” mode, which is the ability to simultaneously work on two different tasks and/or in two different spaces. When asked about preference of ICs in “Sync” mode, 22 subjects preferred to use both ICs, two subjects preferred tablet only, and no subject selected VR only. Different answers were given in the “No-Sync” mode, with nine for both ICs, four for tablet only, and 11 for VR only. In other words, subjects preferred using both ICs with task synchronization, but staying with one IC without it.

The subjects were also asked to give general comments about the HVE level editor. Eleven subjects appreciated the complementary benefits offered by the heterogeneous views and interfaces. They suggested 2D tasks (e.g., painting and menu control), long distance navigation, and large scale manipulation to be performed on the tablet, and 3D tasks (e.g., object selection and scaling), local space locomotion, and small scale adjustment to be performed using immersive VR. Having redundant functionality on both ICs was acknowledged by two subjects, for it granted them freedom to perform the tasks differently in different situations. Lastly, suggestions to improve the HVE level editor were given in the interviews, such as undo and redo (three subjects), ambient sound and sound effects (two subjects), teleport in VR (three subjects), flying in VR (two subjects), showing a virtual tablet in the HMD (one subject), and combining the wand and tablet into a single interface like the Nintendo WiiU controller (one subject).

Video Analysis

To understand how the subjects used the two ICs, we captured videos of the experiment trials from three sources. A web camera was mounted on the ceiling to capture the subject from the top, and screen capture software was installed on the desktop computer and the tablet to capture from both screens. The three streams of video footage for each trial were then merged, timeline-synchronized, and analyzed by the authors. The videos showed that subjects were able to connect the two views in the shared 3D space, and take advantage of both ICs for different tasks. For example, after painting the mountain with the wand, many subjects immediately switched to the tablet, located the river near the mountain, and continued to clean the foliage in it. With task synchronization, the subjects did not need much time to plan such sequences of transitional actions, and were able to execute smoothly. On the other hand, although all subjects eventually adapted to the absence of task synchronization, many of them expressed confusion and awkwardness to repeat actions that had already been done, and some even made a few mistakes when they lost track of the ICs’ individual statuses. The videos also showed that subjects made fewer transitions without task synchronization. They grouped all appropriate tasks for one IC, and completed them before changing to the other IC.

There was also no within-task transition for the cube collecting task in “No Sync” mode. Many subjects chose to stay at the wand, and traveled long distances to carry the cubes to their destinations. This is probably because they had to reselect the same cube on the tablet, which was just why the wand was used in the first place. In contrast, several subjects were able to discover some efficient strategies to leverage both ICs with task synchronization enabled. For example, three subjects completed the cube collecting task quickly by using the tablet to teleport the Hero avatar near a small cube, selecting it with the wand, teleporting with the tablet again near the destination, and dropping the cube. Another interesting approach was taken by two subjects, who positioned the Hero avatar near the destination, and used the wand to drop cubes that have been selected using the tablet from a zoomed-in view.

The "teleport" and "focus" buttons were used a lot in the experiment. Using these two buttons, a subject demonstrated an interesting strategy to speed up multi-scale navigation on the tablet. Instead of panning and zooming in the God camera, the subject teleported his Hero avatar, and tapped the focused button. This allowed him to instantly navigate to an area of interest. However, the "follow" toggle was not used as much, probably because our test bed did not include any "focus + context" task.

Lastly, the video analysis gave us insight about how the interfaces were used for the five test bed tasks. In general, the tablet was mainly used for 2D tasks that needed to be done from different angles, and at large scales, such as painting textures on the terrain, clearing foliage in the rivers, and moving cubes across the VE. In contrast, the wand and HMD were used to edit details of objects in 3D spaces, such as selecting cubes, smoothing terrain surfaces, scaling houses, and planting flowers under trees. These interaction patterns agreed with the subjects' comments in the interview, and clearly indicated the complementary benefits of the two ICs for 3D interaction tasks with diverse requirements.

Discussion

All three hypotheses were confirmed by the user study results. Similar interaction patterns were discovered in the interview feedback and the video analysis, proving that the subjects were able to connect the Hero and God views in the shared virtual space, and learn and use both ICs effectively to perform tasks with diverse and complementary requirements (***H2***). However, the transitions between ICs were much more continuous with task synchronization enabled, as suggested by comparative ratings, user comments in the interview, and video analysis of the experiment trials (***H1***). In comparison, the HVE system without task synchronization was perceived to be confusing, awkward, and inefficient to learn and use in a hybrid way. In essence, the absence of task synchronization broke the hybrid system into two separate tools. Although it was still beneficial to use both ICs for complementary task requirements, subjects tended to avoid transitions as much as possible. The video analysis showed them doing so by dividing the tasks into two groups, and finishing all tasks in one IC before transitioning to a different one. And when some subjects attempted to add more transitional interactions to their workflows, mistakes were made, because they forgot to constantly invest more working memory to keep track of the status of both systems. The synchronization of travel and object selection also enabled and inspired various within-task transition strategies to perform the cube-collecting task efficiently (***H3***). In comparison, these strategies were abandoned when task synchronizations were absent, because subjects had to reselect the cubes in the second IC, which was the reason why it was not used in the first place.

CONCLUSION

To conclude, this paper proposed a novel HVE system to overcome the limitations of traditional immersive VR systems, in task scenarios that involved diverse scales, angles, perspectives, reference frames, or dimensions. The system leveraged the power and rich interactivity of a tablet device to complement the natural yet limiting 3D interfaces in a traditional HMD and wand-based immersive VR setup. The definition of interaction context (IC) was given, and a taxonomy of IC components was presented. Based on research findings in related fields, four coordination mechanisms were proposed to increase the transition continuity between the ICs. And two of them, namely, mutual awareness and task synchronization, were implemented in the current version of the HVE system. Lastly, a user study was conducted based on five level-editing tasks, to validate the benefits of multiple ICs, and compare the transition experience with and without task synchronization enabled. The study results confirmed that complex HVE systems can be learnt and used to perform diverse 3D tasks efficiently, and suggested that task synchronization is necessary to keep continuous and effortless transitions across ICs.

Regarding future work, we are looking to further optimize the transition experience between the ICs through input sharing, and display blend-in techniques, and evaluate the effectiveness of these coordination mechanisms through similar user studies. In addition, we are also interested in applying the same methodology to non-occlusive HMD devices or CAVE based VR systems, as well as experimenting with HVE systems with more than two ICs.

REFERENCES

1. Angus, I. G., and Sowizral H. A. Embedding the 2D interaction metaphor in a real 3D virtual environment. *Proc. IS&T/SPIE's Symposium on Electronic Imaging: Science & Tech. '95*. 282-293.
2. Benko, H., Ishak, E. W., and Feiner, S. Collaborative mixed reality visualization of an archaeological excavation. *Proc. IEEE ISMAR'04*, 132-140.
3. Billinghurst, M., Kato, H., and Poupyrev, I. The MagicBook: a transitional AR interface. *Computers & Graphics, 25*, 5 (2001), 745-753.
4. Bornik, A., Beichel, R., Kruijff, E., Reitinger, B., and Schmalstieg, D. A hybrid user interface for manipulation of volumetric medical data. *Proc. IEEE 3DUI'06*, 29-36.
5. Bowman, D. A., Kruijff, E., LaViola, J. J., and Poupyrev, I. *3D User Interfaces: Theory and Practice*. Addison-Wesley Professional, 2004.

6. Bowman, D. A., Wineman, J., Hodges, L. F., and Allison, D. Designing animal habitats within an immersive VE. *IEEE Computer Graphics and Applications, 18*, 5 (1998), 9-13.
7. Brown, L. and Hua, H. Magic lenses for augmented virtual environments. *IEEE Computer Graphics and Applications, 26*, 4 (2006), 64-73.
8. Bruder, G., Steinicke, F., Valkov, D., and Hinrichs, K. Augmented virtual studio for architectural exploration. *Proc. VRIC'10*, 43-50.
9. Carvalho, F. G., Trevisan, D. G., and Raposo, A. Toward the design of transitional interfaces: an exploratory study on a semi-immersive hybrid user interface. *Virtual Reality, 16*, 4 (2012), 271-288.
10. Churchill, E. F., and Snowdon D. Collaborative virtual environments: an introductory review of issues and systems. *Virtual Reality, 3*, 1 (1998), 3-15.
11. Darken, R. and Durost, R. Mixed-dimension interaction in virtual environments. *Proc. ACM VRST'05*, 38-45.
12. Dubois, E., Nigay, L., and Troccaz, J. Assessing continuity and compatibility in augmented reality systems. *Universal Access in the Information Society, 1*, 4 (2002), 263-273.
13. Feiner, S. and Shamash, A. Hybrid user interfaces: breeding virtually bigger interfaces for physically smaller computers. *Proc. ACM UIST'91*, 9-17.
14. Grasset, R., Dunster, A., and Billinghurst, M. Moving between contexts - a user evaluation of a transitional interface. *Proc. IEEE Artificial Reality and Telexistence'08*, 137-143.
15. Holm, R., Stauder, E., Wagner, R., Priglinger, M., and Volkert, J. A combined immersive and desktop authoring tool for virtual environments. *Proc. IEEE VR'02*, 93-100.
16. Lindeman, R., Sibert, J., and Hahn, J. Towards usable VR: an empirical study of user interfaces for immersive virtual environments. *Proc. ACM CHI'99*, 64-71.
17. Miguel, M. M., Ogawa, T., Kiyokawa, K., and Takemura, H. A PDA-based see-through interface within an immersive environment. *Proc. IEEE Artificial Reality and Telexistence'07*, 113-118.
18. Milgram, P., Takemura, H., Utsumi, A., and Kishino, F. Augmented reality: a class of displays on the reality-virtuality continuum. *Proc. Photonics for Industrial Applications'95*, 282-292.
19. North, C. and Shneiderman, B. Snap-together visualization: a user interface for coordinating visualizations via relational schemata. *Proc. ACM AVI'00*, 128-135.
20. Pierce, J. S., Steams, B. C., and Pausch, R. Voodoo dolls: seamless interaction at multiple scales in virtual environments. *Proc. ACM i3D'99*, 141–145.
21. Plumlee, M. and Ware, C. Integrating multiple 3D views through frame-of-reference interaction. *Proc. IEEE CMV'03*, 34-43.
22. Plumlee, M. and Ware, C. An evaluation of methods for linking 3D views. *Proc. ACM i3D'03*, 193-201.
23. Poupyrev, I., Ichikawa, T., Weghorst, S., and Billinghurst, M. Egocentric object manipulation in virtual environments: empirical evaluation of interaction techniques. *Computer Graphics Forum, 17*, 3 (1998), 41-52.
24. Schmalstieg, D., Encarnacao, M., and Szalavari, Z. Using transparent props for interaction with the virtual table. *Proc. ACM i3D'99*, 147-154.
25. Schmalstieg, D. and Schaufler, G. Sewing worlds together with SEAMS: a mechanism to construct complex virtual environments. *Presence: Teleoperators and Virtual Environments, 8*, 4 (1999), 449-461.
26. Song, P., Goh, W., and Fu, C. WYSIWYF: exploring and annotating volume data with a tangible handheld device. *Proc. ACM CHI'11*, 1333-1342.
27. Steinicke, F., Ropinski, T., Hinrichs, K., and Bruder, G. A multiple view system for modeling building entities. *Proc. IEEE CMV'06*, 69-78.
28. Stoakley, R., Conway, M., and Pausch, R. Virtual reality on a WIM: interactive worlds in miniature. *Proc. ACM CHI'95*, 265-272.
29. Szalavári, Z. and Gervautz, M. The personal interaction panel - a two-handed interface for augmented reality. *Computer Graphics Forum, 16*, 3 (1997), C335-C346.
30. Viega, J., Conway, M. J., Williams, G., and Pausch, R. 3D magic lenses. *Proc. ACM UIST'96*, 51-58.
31. Wang Baldonado, M. Q., Woodruff, A., and Kuchinsky, A. Guidelines for using multiple views in information visualization. *Proc. ACM AVI'00*, 110-119.
32. Watsen, K., Darken, R., and Capps, M. A handheld computer as an interaction device to a virtual environment. *International Immersive Projection Technology Workshop*, 1999.
33. Wilkes, C.B., Tilden, D., and Bowman, D. A. 3D user interfaces using tracked multi-touch mobile devices. *Proc. JVRC of ICAT-EGVE-EuroVR'12*, 65-72.
34. Zanbaka, C. A., Lok, B. C., Babu, S. V., Ulinski, A. C., and Hodges, L. F. Comparison of path visualizations and cognitive measures relative to travel technique in a virtual environment. *IEEE Transactions on Visualization and Computer* Graphics*, 11*, 6 (2005), 694-705.