Augmented Viewport: An action at a distance technique for outdoor AR using distant and zoom lens cameras

Thuong N. Hoang and Bruce H. Thomas
Wearable Computer Lab – University of South Australia
{Thuong.Hoang & Bruce.Thomas}@unisa.edu.au

Abstract

In this paper, we describe the Augmented Viewport, a new action at a distance technique for outdoor augmented reality using wearable computers. Augmented viewport is inspired by the virtual viewport window that has been widely used in virtual reality systems. Our technique utilizes a physical camera, which either is placed at a distant location or can zoom closer to a distant location, to provide the physical world information for the viewport window. Using the augmented viewport, the user can perform selection and affine transformations of distant virtual objects, using techniques that are more effective from a close distance, such as ray casting or image plane. We conducted a user study with results showing that the augmented viewport technique enhances the precision of object manipulations, and reduces time and effort.

1. Introduction

This paper presents a new precise manipulation action at a distance interaction technique for outdoor augmented reality (AR) systems using wearable computers. Users frequently require interactions with virtual objects that are not within arm's reach in outdoor augmented environments. Action at a distance (AAAD) is the problem of interacting with virtual objects that are located at a distance within the user's field of view but out of the user's arm reach [1].

In this paper, we present *augmented viewport*, a new technique allowing AAAD to be performed on an outdoor AR system. Our technique is inspired by the virtual viewport concept for VR [2, 3]. A VR viewport offers a view of the world from a distant location through a virtual window. Similarly, augmented viewports contain both virtual and physical information in the form of a window view of the augmented world at a distant location. Our augmented viewport technique utilizes physical cameras to supply information from a distant location. There are three types of cameras for augmented viewports. A *distant camera* is located at a distance within the user's field of view. A z*oom lens camera* shows a close-up view of the distant world from the user's viewpoint. A *remote camera* shows locations that cannot be seen from the user's position, as with AR telepresence systems.

The augmented viewport supports AAAD in outdoor augmented reality. It enables selection and affine transformation on distant virtual objects by using effective close body interaction techniques, such as ray casting and image plane techniques [1]. By offering a closer and more detailed view of the augmented world, the viewport allows the user to perform precise manipulation operations with reduced errors. These errors are regularly caused by freehand operations or sensor errors that are amplified at distances. The benefits of the technique are validated by a user study.

2. Background

There have been many solutions for the AAAD problem in VR, notably World-in-miniature (WIM) [4], scaled world grab [5], voodoo dolls [6], and image plane techniques [7]. These techniques have similar approaches of bringing the distant world or objects closer to the user. Another approach is to bring the user closer to the objects, as a whole or in parts, such as ray casting, arms/hands extension, and teleportation techniques. Most of these techniques are effective in VR environments where the user's presence is purely virtual and not constrained by the physical environment. The user is able to perform such actions as flying, extending hands, or teleporting. However, in AR systems where the user has both virtual and physical presence, such abilities are not supported. WIM in outdoor AR [8] requires the modeling of large and dynamic outdoor environments, which is time-consuming. AR WIM in general also places the virtual world out of the context with the physical world, diminishing the benefits of AR visualizations.

The virtual viewport VR technique offers the user a view of the world at a distance through a virtual window, as if they were standing in a remote location. Manipulation of distant objects can be performed

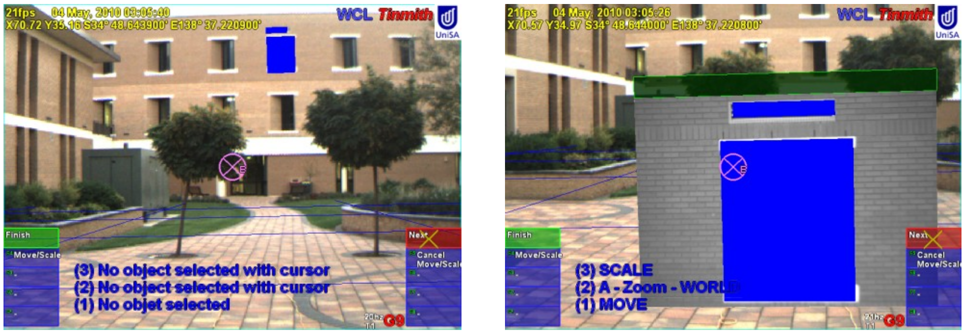

Figure 1: The physical window and virtual objects are shown, (left) without and with (right) the viewport

through the viewports using within arm's reach interaction techniques [3]. SEAM [2] is a mechanism that uses virtual viewports to weave multiple virtual environments together. SEAM utilizes the door/window metaphor to offer instant teleportation without disruption to the user's movement, orientation or the perceived continuity of the virtual world. Through-the-lens techniques [9] implement a similar metaphor for navigation and object manipulation in VR. Multiple viewports can be used to interact with objects from two separate distant locations [3].

There has been limited research on AAAD techniques for AR. Ray casting and the AR working planes techniques [1] enable affine transformation of virtual objects via their 2D projections on a user-specified image plane. These do not support manipulation along the normal axis of the image plane.

3. Augmented viewport

We propose the application of the virtual viewport concept as a solution to AAAD for outdoor AR, called the *augmented viewport* technique. An augmented viewport provides a window for a closer view of the augmented physical world, see Figure 1. It takes the form of a virtual window in which both the physical and virtual views from the distant location are rendered. The distant physical world information is provided in real-time by the viewport camera, which is either a *distant*, *remote*, or *zoom lens* camera, as previously described. Virtual renderings are taken from a viewpoint placed at the same location as the distant or remote camera or with the same zooming ratio as the zoom lens camera. By combining the two sources, the augmented viewport allows the user to see and interact with a remote part of the augmented physical world.

There are three placements of the augmented viewport in the user's view that use different relative coordinate systems. *World relative* places the viewport at a fixed location in the physical world. *Body relative* keeps the viewport always at a fixed distance from the user's body, while *head relative* fixes the viewport to the user's head position and orientation. We deployed our augmented viewport on the Tinmith outdoor video see-through AR system [10]. Augmented viewports are implemented using the Tinmith scenegraph and the OpenGL stencil buffer.

The AAAD technique allows the user to select and manipulate virtual objects more precisely at a distance using image plane techniques [1]. The augmented viewport technique improves precision in AAAD by providing a closer and more detailed view of the remote objects and their physical remote surroundings. The technique reduces freehand and sensor errors that are amplified through the distance. Figure 1 compares the views of the distant location with and without the viewport. The rightmost figure depicts the augmented viewport providing a more detailed view of the distant location.

The augmented viewport removes the need for the user to move physically closer to distant virtual objects and maintains a similar precision when performing manipulation operations. Therefore, it is more efficient and effective. With the telepresence mode of *remote cameras*, the augmented viewport provides the user with a view that is otherwise impossible from their current location, due to physical constraints. AR X-ray vision offers a similar ability [11], but is only limited to visualization without interaction. It could be possible to use static satellite imagery with the augmented viewport, as it has been proven to increase precision in positioning virtual annotations [12].

4. User study

We conducted a user study to evaluate user's preference and performance when using the augmented viewport technique. We had 16 participants (14 males and 2 females), age range 19 – 54. 37.5% had never used a wearable computer or an AR system. The participants were asked to scale and translate virtual boxes to match the sizes and locations of a window and a vent of a distant physical building, using the Tinmith wearable computer. There were two locations: Location **A** at 50m and Location **B** at 30m from the building. The locations were chosen so that the physical window and vent were visually the same size

Table 1: Mean completion time (seconds), Errors in Translation (meters) and Scale (unit)

Cond.	Translation time (s)				Translation errors (m)				Scale time (in s)				Scale errors (in units)			
	Total		From 1st click		Window		Vent		Total		From 1st click		Window		Vent	
	Mean	SD	Mean	SD	X	Y	X	Y	Mean	SD	Mean	SD	X	Y	X	Y
1	20.06	7.10	14.63	8.98	0.40	0.37	0.37	1.39	18.57	6.78	14.93	6.43	0.10	0.11	0.11	0.08
2	17.97	6.01	14.20	6.19	0.48	0.37	0.40	1.37	18.03	5.94	14.75	6.10	0.09	0.06	0.11	0.08
3	17.53	6.43	13.37	6.09	0.30	0.38	0.30	1.39	18.32	5.01	13.95	6.52	0.11	0.12	0.10	0.1
4	24.92	10.41	18.78	10.89	2.01	2.71	1.99	1.16	28.11	10.53	21.15	12.27	0.18	0.22	0.14	0.18
5	43.79	13.98	17.69	11.77	1.67	1.62	1.55	0.91	48.08	13.46	22.15	17.13	0.15	0.26	0.16	0.12
Mean	24.85		15.73		0.97	1.09	0.92	1.24	26.22		17.38		0.13	0.15	0.12	0.11

when viewed from location B as when viewed through the augmented viewport at location A. There were five different conditions to complete the task: location A using the augmented viewport with (1) Head relative placement, (2) Body relative placement, (3) World relative placement, (4) without the augmented viewport, and (5) location B without the augmented viewport. For each condition, the participants were asked to either scale or translate the virtual objects. In total there were ten tasks for each participant to perform. Each participant completed three iterations with a randomized task order and a break in between each task, with prior training.

The augmented viewport technique in the study was implemented using an optical zoom camera. A USB camera with a zoom lens was mounted on a tripod and placed at location A, pointing towards the physical window and vent of the distant physical building. The camera's position and orientation were calibrated and fixed; thus no tracking was required. A trackball mouse was used as an input device. For each task the time required to complete the task was recorded, as the *total time*. The *total time* was calculated from when the researcher activated the task until the participant finished the manipulation. Considering the problem domain of AAAD, for tasks at location B, additional time for the participant to walk from A to B was considered time spent on the task and included in the *total time* record. For comparison, the time period since the participant *first clicked* on the virtual objects until task completion was also recorded, see Table 1. The *first-click time* period did not include walking time for tasks at location B nor initial time in conditions 1-4. A questionnaire was completed at the end of the experiment. *Our hypothesis is that the augmented viewport technique produces greater precision and requires less time to complete the task.*

5. Results

For the *total time* to complete all tasks there was a significant effect ($p < 0.05$) over all five conditions. These were determined by ANOVA analysis, see Table 1. A post-hoc analysis on the *total time* of each task was performed with a pairwise t-Test on the six pairs of conditions [(1,4), (2,4), (3,4), (1,5), (2,5), and (3,5)] with a Bonferroni correction ($\alpha < 0.008$). *For all conditions (1,4), (2,4), (3,4), (1,5), (2,5), and (3,5) there was a significant effect ($p < 0.008$) that the augmented viewport technique reduced the total time to perform translate and scale operations.*

We measured errors as a measurement of distance or scale factor from the correct final transformation. For the translation of the graphical window (in X and Y) and vent (in X) there was a significant difference ($p < 0.05$) over all five conditions by ANOVA analysis. A post-hoc analysis on the translation error measurements was performed with a pairwise t-Test on the six pairs of conditions [(1,4), (2,4), (3,4), (1,5), (2,5), and (3,5)] with a Bonferroni correction ($\alpha < 0.008$). This analysis was only performed on the translation tasks of the graphical window (in X and Y) and vent (in X only). *For all conditions (1,4), (2,4), (3,4), (1,5), (2,5), and (3,5) there was a significant effect ($p < 0.008$) that the AR viewport technique improved translation precision for the graphical window task (in X and Y) and vent task (in X).*

For the scaling of the graphical window (in X and Y) and vent (in Y only) there was a significant difference ($p < 0.05$) over all five conditions by ANOVA analysis, see Table 1. A post-hoc analysis on the scaling error measurements was performed with a pairwise t-Test on the six pairs of conditions [(1,4), (2,4), (3,4), (1,5), (2,5), and (3,5)] with a Bonferroni correction ($\alpha < 0.008$). This analysis was only performed on the scaling tasks of the graphical window (in X and Y) and vent (in Y). *There were a number of significant effects ($p < 0.008$) on the following conditions: the graphical window (in X) between conditions (2,4), the graphical window (in Y) between conditions (2,4) and (2,5), and the vent (in Y) between conditions (1,4), (2,4), and (3,4).* This showed that our technique aided in the scaling tasks at distances.

Our results validate the benefits of the augmented viewport technique. Participants preferred the augmented viewport technique for manipulating objects at a distance (1, 2, and 3, combined in the chart in Figure 2), over manipulating objects at a closer distance (5) and manipulating objects at a distance without augmented viewport (4). The participants were asked to rate how easy the tasks were and how precise they thought their manipulations were. Figure 2 shows mean scores of the responses, on the scale of 0-100%, with 100% meaning Easy, and 0% Hard.

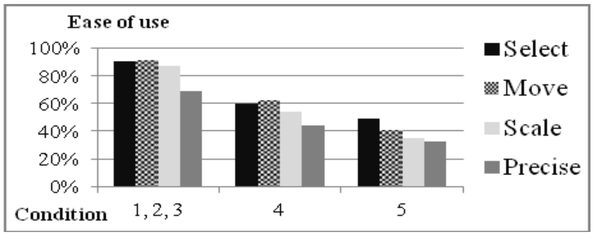

Figure 2: Ease of use preferences

The participants commented that they preferred the augmented viewport most because it offers a larger, magnified viewing area, with a clearer camera image allowing for more precise control. We believe that the optical zoom has a possible advantage over the magnified GPU zoom in terms of pixel clarity. Location B without the viewport is the second preferred condition. The main feedback from the participants at location B is the discomfort of having to look up quite high to complete the task to the point of causing neck strain, considering that the physical window was approximately 10m high and the participant was standing at 30m from the wall. The least preferred condition is at location A without the zoom viewport, because the head mounted camera could not provide enough details of the physical building at such distance. There were a few occasions when the participant could not see the physical vent. There is no definite preference over the placement of the viewport (world, body, and head relative).

The stable mount of the zoom camera does not explain the observed performance improvements. When the user is stationary, for world and body relative placements, the augmented viewport has a fixed location, independent of the user's head orientation and input device. Therefore, both the viewport (condition 1, 2, 3) and the physical building (condition 4, 5) are subject to the same jittering errors of the input device and HMD. The difference is the errors attenuation in conditions 4, 5 over the distance.

It is noted that the implementation of the condition at location A without using the viewport is the working plane technique [1]. Therefore, we have concluded that the augmented viewport technique is a more effective AAAD technique for outdoor augmented reality.

6. Conclusion

We present the augmented viewport as an AAAD technique for outdoor augmented reality using a wearable computer. The augmented viewport provides a window view into the augmented environment at a distance, via the use of physical distant, remote, or zoom lens cameras. The technique allows the manipulation of virtual objects at a distance, utilizing close hand interaction techniques. Through a user study, we have validated that the augmented viewport technique improves precision in selection and manipulation tasks, as well as saves time and effort.

7. Acknowledgements

We would like to thank Ross Smith for his help with the Tinmith backpack, Wayne Piekarski as the developer of Tinmith, and members of the WCL lab for proofreading the paper.

8. References

[1] W. Piekarski and B. H. Thomas, "Augmented reality working planes: a foundation for action and construction at a distance," *ISMAR 2004*, pp. 162-171, 2004.

[2] D. Schmalstieg and G. Schaufler, "Sewing Worlds Together With SEAMs: A Mechanism to Construct Complex Virtual Environments," *Presence: Teleoperators & Virtual Environments*, vol. 8, pp. 449-461, 1999.

[3] K. Hirose, T. Ogawa, K. Kiyokawa, & H. Takemura, "Interactive Reconfiguration Techniques of Reference Frame Hierarchy in the Multi-viewport Interface," in *IEEE 3DUI 2006*, pp. 73-80, 2006.

[4] R. Stoakley, M.J. Conway, & R. Pausch, "Virtual reality on a WIM: interactive worlds in miniature," *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 265-272, 1995.

[5] M. R. Mine, Jr.F.P. Brooks, & C.H. Sequin, "Moving objects in space: exploiting proprioception in virtual-environment interaction," *24th Conference on Computer graphics & interactive techniques*, pp. 19-26, 1997.

[6] J. S. Pierce, B.C. Steams, & R. Pausch, "Voodoo dolls: seamless interaction at multiple scales in virtual environments," *Proceedings of the 1999 symposium on Interactive 3D graphics*, 1999.

[7] J. S. Pierce, A.S. Forsberg, M.J. Conway, S. Hong, R.C. Zeleznik, & M.R. Mine, "Image plane interaction techniques in 3D immersive environments," *Proceedings of the 1997 symposium on Interactive 3D graphics*, 1997.

[8] B. Bell, et al., "An annotated situation-awareness aid for augmented reality," *15th ACM UIST*, 2002.

[9] S. L. Stoev and D. Schmalstieg, "Application and taxonomy of through-the-lens techniques," in *ACM VRST 2002*, pp. 57-64, 2002.

[10] W. Piekarski and B. H. Thomas, "Tinmith-evo5 - an architecture for supporting mobile augmented reality environments," in *ISMAR 2001*, pp. 177-178, 2001.

[11] B. Avery, B.H. Thomas, & W. Pierkarski., "User evaluation of see-through vision for mobile outdoor augmented reality," the *IEEE/ACM ISMAR*, 2008.

[12] J. Wither, S. DiVerdi, & T. Hollerer, "Using aerial photographs for improved mobile AR annotation," in *ISMAR 2006,* pp. 159-162, 2006.