Integrated Platform for an Augmented Environment with Heterogeneous Multimodal Displays

Jaedong Lee, Sangyong Lee, and Gerard Jounghyun Kim

Digital Experience Laboratory
Korea University, Seoul, Korea
`{jdlee,xyleez,gjkim}@korea.ac.kr`

Abstract. With the recent advances and ubiquity of various display systems, one may configure an augmented space with a variety of display systems, such as 3D monitors, projectors, mobile devices, holographic displays, and even non-visual displays such as speakers and haptic devices. In this paper, we present a software support platform for representing and executing a dynamic augmented 3D scene with heterogeneous display systems. We extend the conventional scene graph so that a variety of modal display rendering (aside from just visual projection) can be supported. The execution environment supports multi-threading of the rendering processes for the multiple display systems and their synchronization. As multiple and heterogeneous displays, in effect representing a particular set objects in the augmented environment, are scattered in the environment, additional perception based spatial calibration method is also proposed.

Keywords: Augmented space, Extended scene graph, Multiple displays, Calibration, Floating image display.

1 Introduction

An augmented environment refers to a physical 3D environment spatially and naturally registered with virtual objects. Being "natural" means that the virtual objects are perceived to mix in with the physical environment seamlessly and felt as everyday objects. Tight spatial registration means virtual objects situated in the right location and pose, and this would be one requirement for naturalness. With the recent advances and ubiquity of various display systems, realization of such "naturally" augmented environments has become viable. One may configure an augmented space with a variety of display systems, such as 3D monitors, projectors, mobile devices, holographic displays, and even non-visual displays such as speakers and haptic devices (Figure 1). In this paper, we present a software support platform for representing and executing a dynamic augmented 3D scene with heterogeneous display systems. We extend the conventional scene graph so that a variety of modal display rendering (aside from just visual projection) can be supported. The execution environment supports multi-threading of the rendering processes for the multiple display systems and their synchronization. As multiple and heterogeneous displays scattered in the environment does not easily lend themselves to the conventional, e.g. computer vision, based calibration process, a perception based spatial calibration method is proposed.

R. Shumaker (Ed.): VAMR/HCII 2013, Part I, LNCS 8021, pp. 73–78, 2013.

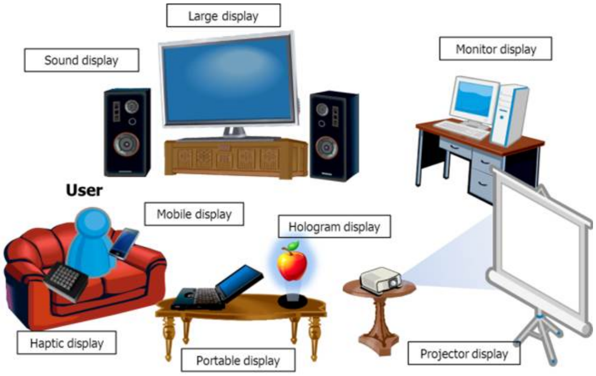

Fig. 1. An augmented environment with multiple heterogeneous displays

2 Related Work

Virtual reality environments often require their presentations through large displays. For that purpose, researchers have proposed methods for distributed rendering and tiled display [1, 2, 3]. However, these systems typically do not address the use of different types of display systems, nor do they consider displays systems that are dispersed in the environment through which augmentation environment/object can be viewed. One notable exception is the work by Yang et al. who proposed a layered display system [4] and switching between heterogeneous display systems in the environment as the users moves around in it for an optimized viewing condition. In our work, we focus on a comprehensive (in terms of types of display systems supported, synchronization, and calibration) software platform for handling multiple heterogeneous displays.

3 Scene Graph Extension

3.1 Object Types / Display Parameters

We took a conventional scene graph data structure, and extended it so that it can represent different types of objects and specify associated display systems and required parameters. For instance, an object type may be a real physical object with its geometry and attributes captured by sensors, reconstruction methods and even direct measurements. An object may be purely virtual (augmented into the environment) and designated to be presented visually through a holographic device and with tactile feedback through a vibration device. Note that for a given object and its modality, the developer may wish to specify a particular display rendering algorithm. Node types and attributes have been revised support specification of such information and parameters so that when the scene graph is processed, it can be used for proper rendering and synchronization. With the extended node sets, the developer can specify and design the augmented environment with more ease and without having to worry about low level details.

3.2 Synchronization

There are two main objectives for synchronization. One is for among image frames of different visual displays, and the other for among different modality output corresponding to a particular event. For example, when virtual ball is dropped, it may be rendered visually, aurally and with force feedback, and all the modal output must occur with minimal temporal delay to be felt as one event.

For the former, similarly to the techniques employed by tiled display systems, we use software “gen-lock” to synchronize image frames among within 2~3 frame difference [5, 6]. Such degree of difference is usually regarded visually unnoticeable. In our case, the slowest rendering node serves as the reference to all other rendering nodes for synchronization (see Figure 2). While frame coherence is guaranteed, the temporal coherence may suffer if there were many rendering nodes to which network messages must be sent. However, in normal situations, there would not be so many rendering nodes (e.g. less than 20~ 30). What is important is that upon an event its multimodal after effects are rendered simultaneously. When different visual displays are rendered on separate graphics card, the hardware gen-lock can be used [7].

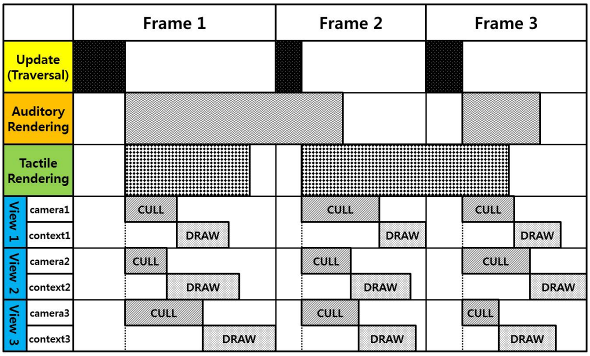

Fig. 2. Three rendering nodes synchronized. For example, the visual rendering of Frame 2 only starts after view 3 has finished rendering Frame 1. In the meantime, the aural and tactile rendering starts off upon traversing the scene graph and invoking the corresponding call back threads, which may sometimes continue over the visual frame boundary. The aural and tactile rendering, however, does not occur as frequently as the visual (which must occur at least 15~20 times per second).

As for the latter objective, our extension provides a protocol for different modal displays to refer to a common event specification and synchronize each multimodal output thread around it by a call-back mechanism. For example, the aural and tactile rendering starts off after traversing the scene graph by restarting the corresponding the call-back threads, which may sometimes continue over the visual frame boundary.

Practically, this is not problematic as the aural and tactile events and their rendering, however, do not occur as frequently as the visual (which must occur at least 15~20 times per second).

A test system using the proposed platform has been set up and has been evaluated in terms of the overall frame rate (vs. number of displays sustainable) and temporal synchronization error. Figure 3 shows the test system in which three (and more) different displays are used, a 3D stereoscopic TV (showing the airplane), floating image device (showing the missile) and vibro-tactile device. The objects are spatially registered and synchronized such that when interacted upon, the missile would render a tactile feedback, fly and shoot down the airplane. Our have shown acceptable performance (up to ~20 fps) for supporting up to 12 different displays with each node handling up to more than 120,000 polygons, and exhibiting inter-node frame coherence, and virtually no temporal delay. Such a performance level is deemed sufficient to support and implement a small augmented room.

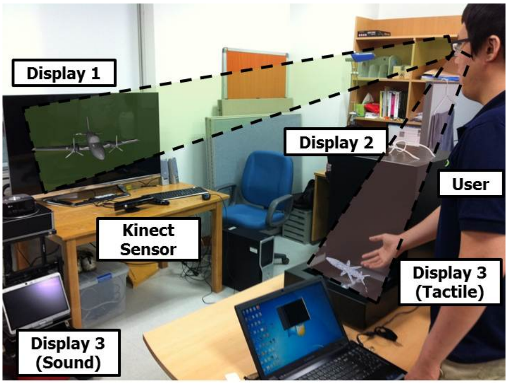

Fig. 3. Test environment consisting of three (and more) different displays (3D TV, floating image display, vibro-tactile device, etc.) in an augmented environment.

4 Calibration

As virtual objects are situated and registered to the physical world through associated display systems, there needs to be a method for these display system to be calibrated according to the units of the physical world and spatially registered in the whole environment. While standard calibration methods and warping methods exist for monitors and projectors, since the augmented space also use other types of displays, e.g. holographic floating image device, we devised a new calibration method, since its display mechanism is different from that of the monitors/projectors. The main difference is that its optical system floats an image in 3D space and it is very difficult to establish the correspondence between the ground truth and actual display points due to the geometry of the display apparatus and other restrictions. For example, the floating image cannot be seen from the front direction due to the location of the reflecting mirrors and other optical elements. The ground truth object has to be suspended in 3D space in a "skeletal" form not to occlude any reflecting lights in forming the floating image (see Figure 4). Installation of such a grid structure object is not always possible. Then, one could take a picture, from a known view point location, of the floating image overlaid on the ground truth skeletal object, and apply image processing techniques to extract the "points" and a set of correspondence match. However, again this is problematic because the imagery is usually not conductive to corner point extraction (e.g. the interior of the display device is dark and the skeletal grid points are barely visible).

Instead we rely on a human to judge whether the floating grid points coincide with those of the ground truth grid. Any perceptual differences are made coincident by adjusting the corresponding points in the virtual space. Such differences (between the virtual points of the ground truth and the virtual points of the adjusted ones) are recorded and later applied for image correction by interpolating the adjustment values. Figure 4 also shows the results of applying the calibration. Note that such calibration is inevitably view dependent and user dependent because the optical system for the floating imagery is complicated resulting in different distortion depending on the view point and because we rely on human judgments (particularly in depth assessments). Thus, ideally, the calibration must be performed at different nominal view points and for customized to the individual user.

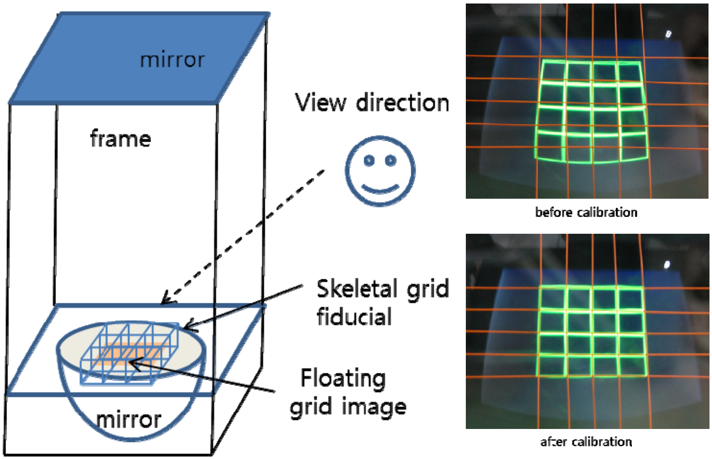

Fig. 4. Calibrating the floating image device used in our augmented space. A skeletal grid representing the ground truth is attached and suspended in the middle of the display space above the spherical mirror. A floating grid image is compared to the ground truth and made coincident perceptually from a given view point by adjusting the corresponding points in the virtual space. The adjustments are recorded and applied to other objects through interpolation.

5 Conclusion

In this paper, we have presented a software support platform for representing and executing a dynamic augmented 3D scene with heterogeneous display systems.

While more performance testing and optimization is required, it offers a convenient software layer abstraction for realizing augmented environments. We hope that such an infrastructure will contribute to proliferate the use of augmented environments for various applications to such as entertainment and education.

Acknowledgement. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No 2012-0009232)

References

1. Li, C., Lin, H., Shi, J.: A survey of multi-projector tiled display wall construction. In: Multi-Agent Security and Survivability, pp. 452–455 (2004)
2. Ni, T., Schmidt, G.S., Staadt, O.G., Livingston, M.A., Ball, R., May, R.: A survey of large high-resolution display technologies, techniques, and applications. In: Virtual Reality Conference, pp. 223–236 (2006)
3. Krumbholz, C., Leigh, J., Johnson, A., Renambot, L., Kooima, R.: Lambda table: high resolution tiled display table for interacting with large visualizations. In: 5th Workshop on Advanced Collaborative Environments (2005)
4. Yang, U.Y., Kim, H.M., Kim, J.H.: Expandable 3D Stereoscopic Display System. Korea Patent Applied, No. 2012-0018690 (2012)
5. Van Der Schaaf, T., Renambot, L., Germans, D., Spoelder, H., Bal, H.: Retained mode parallel rendering for scalable tiled displays. In: Immersive Projection Technologies Symposium (2002)
6. Igehy, H., Stoll, G., Hanrahan, P.: The design of a parallel graphics interface. In: 25th Annual Conference on Computer Graphics and Interactive Techniques, pp. 141–150. ACM (1998)
7. Nvidia, Inc. Genlock, `http://www.nvidia.com/object/IO_10793.html`