Rubberneck Navigation: Manipulating Frames of Reference for 3D Navigation with 2D Input Devices

Desney S. Tan[1], Mathew J. Conway[3], Ken Hinckley[3], Takeo Igarashi[2], Dennis R. Proffitt[4], George G. Robertson[3], and Randy Pausch[1]

School of Computer Science[1]
Carnegie Mellon University
2306 Wean Hall
Pittsburgh PA 15213-3891
{desney, pausch}@cs.cmu.edu

Department of Information Engineering[2]
University of Tokyo
7-3-1 Hongo Bunkyo-ku
Tokyo 113-8656, JAPAN
takeo@mtl.t.u-tokyo.ac.jp

Microsoft Research[3]
One Microsoft Way
Redmond WA, 98052-6399
{mconway, kenh, ggr}@microsoft.com

Department of Psychology[4]
University of Virginia
Gilmer Hall
Charlottesville VA 22903-2477
drp@virginia.edu

ABSTRACT

We present a novel technique for allowing users to control both viewpoint motion and orientation using a 2D input device with a button to provide cognitively simple and unobtrusive navigation in desktop 3D virtual environments. In this task, the user must control the three frames of reference—environment, body, and head—with a single 2D input device. The underlying observation to our solution is the fact that users' control of their frames of reference may be separately, but transparently, constrained in the two dissociated modes of use: wayfinding and travel. When users are stationary, we couple the head to the body and allow them to control the orientation of their entire being and to specify a 3D path to control future viewpoint motion. While in motion along this path, we decouple control of the head from the body, giving users the flexibility to look around the environment while the body continues to move along the specified path, a process we term *rubbernecking* (Figure 1). With a button click, users may toggle between modes and may re-specify paths at any time.

KEYWORDS: Desktop virtual reality, virtual environments, navigation, wayfinding, travel, frames of reference.

INTRODUCTION

The improvement of graphics computations and network capabilities presents the ability to create large virtual environments (VEs). Although these advancements increase our potential to migrate traditional 2D interfaces to 3D virtual workspaces, new human computer interaction techniques are required to tap the human skills acquired from working in the real world. In particular, navigation in immersive and non-immersive VEs has been a challenging problem. Since most VEs encompass more space than can be viewed from a single vantage point, users have to be able to move efficiently within the environment in order to obtain different views of the scene and establish a sense of presence within the 3D space.

Figure 1: Traditional desktop navigation without viewpoint control (*left*); Rubberneck navigation with viewpoint control (*right*)

3D navigation metaphors may be broken into two categories, immersive and non-immersive. Most immersive navigation metaphors include some notion of the body and head as separate entities. Successful navigation metaphors have recognized that it is important for a user to be able to control movement (of the body) as well as exploration (with the head). In contrast, most non-immersive desktop navigation metaphors are based on some notion of camera control (Figure 2). These metaphors are built on gaze-directed steering and do not permit exploring viewpoints other than in the direction of motion. Because of the static mapping of input to control, looking around while moving, or *rubbernecking*, is not possible.

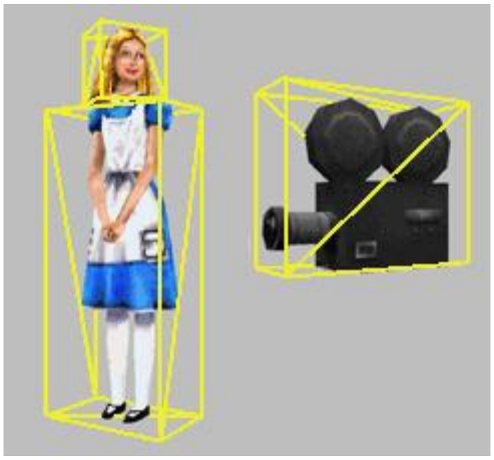

Figure 2: Immersive metaphor - head and body separation; Non-immersive metaphor - single camera

Mapping 2D input devices to effectively control 3D navigation is a difficult problem. Frequently, there is a poor match between the input device, goals of the navigation activity, and the skills of the user. In the real world, people essentially have 9 degrees of freedom (DOF): 3 for the translation of the body, 3 for the orientation of the body, and 3 for the orientation of the head. It can be shown that in navigation tasks, people do not roll or pitch their bodies nor do they, in most cases, roll their heads. Thus we are left with 6 DOF. It requires a large amount of cognition to coordinate between perception and motor skills in order to map the 2 DOF input to the 6 DOF required for movement and exploration. Users often get lost, frustrated, and experience poor information acquisition with such interfaces because their concentration is entirely consumed by the navigation task.

In this paper we consider the use of a 2D input (trackball, mouse, touchpad, pen and stylus) to specify and control movement along a path and to control the viewpoints into the desktop VE. The principles we investigate are based on the cognitive issues surrounding VE navigation, especially with constraining and manipulating the number of degrees of freedom a user has control over. We present a simple interface that provides all the DOF required for the navigation task without placing unnecessary constraints on the user. We show how it is possible to constrain the 9 DOF to at most 4 DOF and then manipulate the coupling of head to body in the different modes to provide a transparent mapping from a 2D input device.

We recognize that wayfinding and travel are dissociated tasks. When wayfinding, users should be able to specify their intended path. When traveling however, users should be allowed to look around to view not only target objects but arbitrary ones within environment. Although these two tasks are related, users usually perform them sequentially, quickly switching back and forth between wayfinding and travel modes. With these principles in mind, we design a desktop navigation metaphor to include the notion of body and head as separate frames of reference. We allow the user to specify paths in the wayfinding mode while the head is coupled to the body, and to look around the environment in the travel mode while the head is decoupled from the moving body. Our technique finds the middle ground between the triviality of an author-designed single camera animation path and the confusing excess freedom of common unconstrained control paradigms.

RELATED WORK

Various implementations and studies for viewpoint motion control in non-immersive 3D environments have been described. Mackinlay et al. describe input devices as transducers of any combination of degrees of freedom and provide complex mouse-based virtual input for simulated 3D egocentric motion control [14]. In a related paper, they present an intelligent target oriented technique movement with 2D and higher dimensional input [15]. Strommen compares three methods, Click go/Click Stop, Hold and go, and Slide and go as mouse-based interfaces to point-of-view navigation [22]. Ware and Slipp assess the usability of different velocity control interfaces for viewpoint control in 3D graphical environments [25]. Chen et al. explore 3D rotation of objects with 2D input devices [4]. More recently, Hanson and Wernert focus on using 2 DOF input devices to move through 3D environments with designer imposed constrains [11]. They propose the “2D virtual sidewalk” specified by a customizable algorithm to compute viewing parameters from the user. Zeleznik and Forsberg construct gestural camera controls for birds-eye view manipulation [26].

The Tesla BattleTech System by Virtual World Entertainment allows the user to move in one direction and look in another using a combination of foot pedals and joystick and button inputs [1]. Virtual Worlds admits a “steep learning curve” for this system. Activision’s MechWarrior Combat Simulation Series provide a similar feature controlled with special input devices, such as the Microsoft Sidewinder joystick, which has the ability to rotate [17]. Id Software’s Quake utilizes a combination of keyboard commands for body motion and mouse controls for head orientation [20]. Galyean presents a “river analogy” to guide users’ motion in a head-tracked immersive environment while allowing the head to look around [10]. Apart from navigation, but similar to the principles we apply, Pierce et al. separate the head and body and use glances, or preset views, to locate associated toolspaces [19]. Igarashi et al. describe a novel way for projecting paths drawn in 2D to the 3D environment [12]. A large part of our work has been inspired by and may be considered an extension to their path drawing techniques. Additionally, Cohen et al. specify non-planar 3D curves to address the problem of flying through space [4].

A large body of literature documents different studies in 3D navigation. These studies include work aiming to evaluate different navigation techniques [2, 3, 19, 21] as well as work exploring spatial orientation and wayfinding [6, 7, 8, 9] and locomotion or travel [15, 24] using various metaphors. In this paper, we extend the body of work to explore desktop 3D navigation using the immersive navigation metaphors of separated body and head.

BACKGROUND

One of the most basic and universal interactions within VEs is navigation. We break the navigation task into two components: wayfinding and travel. Wayfinding is defined as the cognitive process of determining a path based on visual cues, knowledge of the environment, and other aids such as maps. When wayfinding, therefore, users must be able to specify the path on which they wish to move. Travel is defined as the control of a user's viewpoint motion in 3D. The primary role of travel is to transport users from one place to another to allow them to perform some more useful task. Along this path, users must be able to control their viewpoints in order to look around the environment.

In most desktop 3D navigation systems, users control a single entity both when wayfinding and traveling. In these systems, viewpoint orientations are restricted to facing a direction tangent to the path while traveling. This, we assert, is overly restrictive for effective navigation tasks. Using these metaphors, for example, it is difficult to move parallel to a wall while constantly inspecting it or to move around an object while constantly looking at it. Users defining a path must consider not only the actual motion intended but also the orientation of the viewpoint at each point along this path. While various systems have implemented special purpose solutions for this [1, 16, 20], we define a general purpose system which allows the decoupling of head and body, allowing users to *rubberneck*, or look around even when moving. Rubbernecking is an American idiom that describes the action of turning to look at something while moving by it; for example, drivers rubberneck when they look out their car windows at an accident in the next lane while driving by.

Descriptions of relations between objects in any environment are related to reference frames that are tightly coupled to either the user or the environment that the user is in. There exist 3 distinct frames of reference for any navigation task: environment, body, and head. The environmental frame of reference defines the user's position in the environment. This frame of reference implies 3 DOF of translation. The body is a 3D egocentric frame of reference that defines the user's orientation in the environment. This frame of reference implies 3 DOF of rotation of the body. The head, on the other hand, is a 2D visual frame of reference projected forward in the field of view that has 3 DOF of rotation. The user thus has a total of 9 DOF to control in the real world.

REDUCING DEGREES OF FREEDOM

Most of the 9 DOF are not critical, or even essential, in navigation tasks. Here we will describe why we are able to remove at least 5 DOF without significantly affecting the user's freedom of movement or ability to explore the environment.

Only yaw and forward translation are required for the body to follow any given planar path that is parallel to its plane of motion. Along this path, the body constantly moves forward while turning to maintain an orientation tangent to the path at the current point. This motion is similar to driving a car along a road. Thus, assuming we do not allow flying, we may remove left/right and up/down translation of the body. Users typically have no need to control the pitch or roll of the body while wayfinding or traveling; nor do they care to control the roll of the head in most cases. In fact, direct manipulation of these extraneous DOF sometimes leads to more difficult controls. We remove users' control over them. Making these simplifications, we have 4 DOF remaining: forward/backward translation, yaw of the body, and pitch and yaw of the head (Table 1).

	Wayfinding	Travel
Forward/Back	Stationary (unused)	Forward at constant speed
Left/Right	Replaced with body yaw/forward	
Up/Down	Constrained to planar walk	
Body Yaw	**Left/Right**	Tangent to path at all points
Body Pitch	Removed extraneous DOF	
Body Roll	Removed extraneous DOF	
Head Yaw	Attached to body	**Left/Right**
Head Pitch	**Forward/ Backward (optional)**	**Forward/ Backward (optional)**
Head Roll	Removed extraneous DOF	

Table 1: Degrees of Freedom constrained (shaded), implicitly controlled (normal), and user controlled (bold)

For certain applications, it may be appropriate to further remove control of pitch of the head, leaving 3 DOF. This remains a design question to be answered on a task specific basis. For the purpose of simplicity, we will not include pitch of the head in discussions hereafter; this DOF may be mapped to the forward/backward motion of the 2D interface and included trivially. Alternatively, the forward/backward input may be used to control speed of motion. Because of the varying amount of noise between x and y inputs (in different input devices and with different users), we could include a “dead-zone” around the center of interaction to ignore small deflections both in x and y.

MANIPULATING FRAMES OF REFERENCE

A mapping of 2D input to control 3 DOF is still a difficult problem. We use cognitive principles to create an intuitive and transparent mapping that may then be used for efficient navigation. As described earlier, navigation may be broken into two components, wayfinding and travel. We make use of this distinction to separate the tasks into two modes and create our mapping (Table 2).

	Wayfinding	Travel
Environment	Set desired path	Constant motion along path in fixed orientation
Body	**Rotate in tandem, define path to travel**	
Head		**Rotate independently**

Table 2: Coupling/decoupling user controlled (bold) frames of reference

When users are stationary, they are considered to be in the wayfinding mode. In this mode the body and head are coupled and treated as a single control entity. The DOF that the user may control at this point is therefore the yaw of the entire body (Figure 3). This parallels traditional gaze directed steering techniques, where users expect to move in the direction of their gaze. For this purpose, we position targets within the environment that will be highlighted when the user is looking directly at them. Highlighting a target implicitly defines a linear path to that target. Given environmental information, we can extend this to plan paths around obstacles. In addition, we depart slightly from the linear path selection method by allowing users to specify an arbitrary path on which they wish to travel using Igarashi’s Path Drawing technique [12].

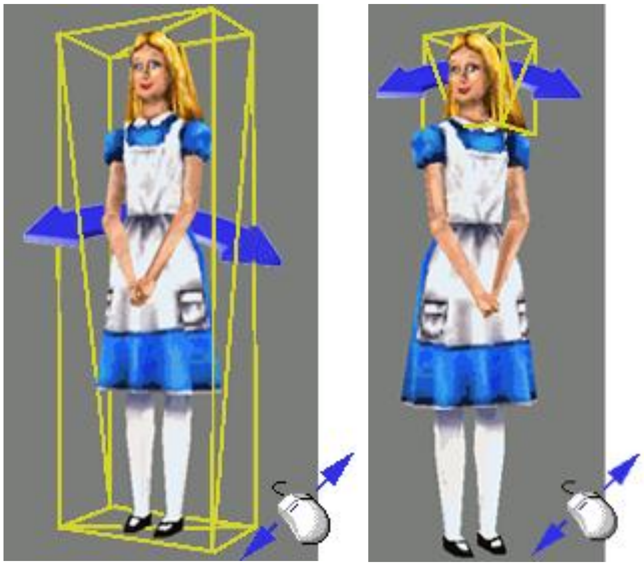

Figure 3: Input controls: yaw of head and body in wayfinding mode (*left*); yaw of head only in travel mode (*right*).

When users hold the input button down and drag in the wayfinding mode, their viewpoint control changes to one of path drawing. In this technique, the system projects the stroke drawn on the 2D screen onto the scene and constructs the intended motion path (Figure 4). With the detailed representation and structure of the virtual space, the system is able to extrapolate such features as obstacle avoidance and moving on uneven surfaces (such as slopes). When users release the button, indicating that the path is complete, they move into travel mode, moving the body along the path. In keeping with the target-oriented metaphor and in order to provide a natural visual flow field, we use a gentle acceleration both at the beginning and end of the path.

In travel mode, control of the body is decoupled from that of the head and the user now has full control of the yaw of the head (Figure 3). Users therefore have full freedom to look around the environment while the body remains oriented in a direction parallel to the tangent of the path at any given point. The body continues to move straight ahead along this path until the user clicks the button or until it reaches its destination. The user may decide to click the input button at any point along the path in order to stop the motion and switch back to wayfinding mode, either to inspect an object along the path more carefully or to define a new path. Because the path drawing technique is meant to be a lightweight tool, switching back into wayfinding mode by clicking the button destroys the current path. When switching back to wayfinding mode, the body is transparently returned to the orientation of the head and is re-coupled to the head.

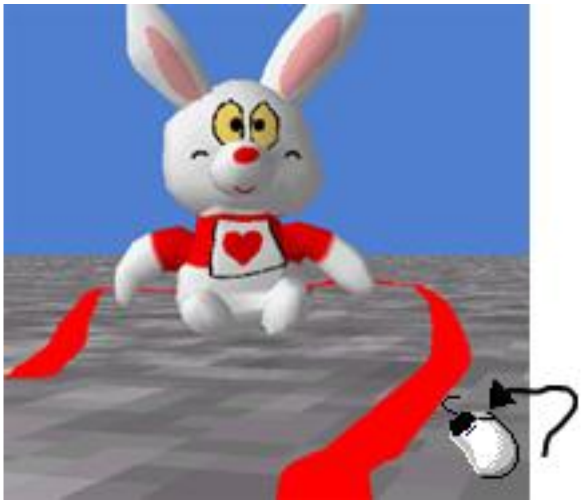

Figure 4: Path drawn by projecting 2D input onto the ground plane.

MERITS AND LIMITATIONS

This technique has been used in a study of Functional Magnetic Resonance Imaging (fMRI) at the University of Virginia. The fMRI allows us to see what part of the brain is activated when we think, feel, move, or experience certain emotions. Clinical fMRI scanners require patients to lie in a narrow, tunnel like openings and perform certain navigation tasks in VR. In this study, the target users are patients with mild brain damage; not only were most users fearful to some extent of the imaging process, but none had training with the input device. Users tested did not have brain damage and were trained for about five minutes before using the system. Most of the users were able to complete tasks given with the interface, giving us great confidence that we have indeed found a new worthwhile human computer interaction technique, especially given the conditions under which they had to perform the task.

The methods for constraining DOF and coupling and decoupling frames of reference provide an abstraction for mapping 2D input to control desktop 3D navigation. The techniques described here limit users' freedom of control just enough to prevent cognitive overload, but not so much as to disturb the feeling of exploration and discovery appropriate to the navigation task. Users remain largely unaware of the reduced DOF. For example, the path drawing technique in conjunction with the coupling techniques gives users a sense of controlling the 3 translation DOF as well as the body yaw DOF, when they actually only have to control two of these. This fact coupled with the fact that we remove unnecessary DOF, makes navigation a low cognition and intuitive task to perform.

In the current implementation, we have not explored the concept of "flying" through space. The current version of the path drawing technique requires a surface to map onto and does not support mapping 2D strokes into free space. Although it is not clear that this would enhance the navigation task, this is an area that remains to be explored. Also, the fact that users have to consciously stop and return to wayfinding mode in order to change paths may be slightly disruptive. We could have implemented a technique whereby the user could move into path drawing while in motion. However, we suspect that this would be a difficult task for novices to master.

FUTURE WORK

We have presented a technique that constrains DOF required for navigation tasks as well as effectively manipulates and abstracts the 3 frames of reference involved in any navigation task. This is a compelling idea that we hope will stimulate thought and further research. Extensions to the current system may include speed control and on-the-fly (while traveling) path drawing and selection. Another interesting direction of this work would be to explore the creation of new input devices that are based on the concept of representing the 3 frames of reference.

Head mounted display tracking may be coupled with image plane interaction techniques as an extension of our work in immersive environments. Here, users could specify targets or draw paths with their fingers or hands and turn their heads to look around while their bodies follow this path.

CONCLUSIONS

We have shown that the 9 DOF that a user normally has may be reduced to at most 4 DOF for desktop 3D navigation tasks. We have further devised a transparent coupling and decoupling of the 3 frames of reference – environment, body, and head – in order to map the 2D input to control the 3 or 4 DOF that we give the user. In the wayfinding (stationary) mode, we couple the head to body and give the user control of the orientation and implicitly of the direction of motion. From this mode a user may either specify a path to take or switch to travel mode, in which control of the body is decoupled from the head. In travel (moving) mode, the user has control of the orientation of the head, thus allowing for *rubbernecking*, or full viewpoint orientation control separate from the actual motion of the body. This technique takes advantage of cognitive principles to create a method for simple, intuitive control with a 2D input device for navigation tasks in desktop virtual environments.

REFERENCES

1. BattleTech: No Guts No Galaxy. Visit http://www.virtualworld.com/ for more information.

2. Bowman, D. Koller, D., and Hodges, L. A Methodology for the Evaluation of Travel Techniques for Immersive Virtual Environments. *Virtual Reality: Research, Development, and Applications*, vol. 3, no. 2, 1998, pp. 120-131.

3. Bowman, D., Koller, D., and Hodges, L. Travel in Immersive Virtual Environments: An Evaluation of Viewpoint Motion Control Techniques. *Proceedings of the Virtual Reality Annual International Symposium (VRAIS)*, 1997, pp. 45-52.

4. Chen, Michael, Mountford, Joy S., Sellen, Abigail. A Study in Interactive 3-D Rotation Using 2-D Control Devices. *Proceedings of the 15th annual ACM Conference on Computer Graphics* , 1988, pp. 121-129.

5. Cohen, Jonathan M., Markosian, Lee, Zeleznik, Robert C., Hughes, John F. An Interface for Sketching 3D Curves. In *Proceedings of the 1999 symposium on Interactive 3D Graphics*, 1999, pp. 17-21.

6. Darken, R.P., Allard, T., & Achille, L. Spatial Orientation and Wayfinding in Large-Scale Virtual Spaces: An Introduction. *Presence: Teleoperators and Virtual Environments*, 7(2), 1998, pp. 101-107.

7. Darken, R.P., & Sibert, J.L. Wayfinding Strategies and Behaviors in Large Virtual Worlds. *Proceedings of ACM SIGCHI 96*, 1996, pp. 142-149.

8. Fukatsu, Shinji, Kitamura, Yoshifumi, Masaki, Toshihiro, and Kishino, Fumio. Intuitive control of “Bird's Eye” Overview Images for Navigation in an Enormous Virtual Environment. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* 1998 , pp. 67 – 76.

9. Furnas, George W. Effective view navigation. *Conference proceedings on Human Factors in Computing Systems* , 1997, pp. 367-374.

10. Galyean, Tinsley A. Guided Navigation of Virtual Environments. *Proceedings of the 1995 symposium on Interactive 3D graphics*, 1995, pp. 103.

11. Hanson, Andrew J., and Wernert, Eric A. Constrained 3D navigation with 2D controllers. *Proceedings of the conference on Visualization '97*, 1997, pp. 175.

12. Igarashi, Takeo, Kadobayashi, Rieko, Mase, Kenji, and Tanaka, Hidehiko, Path Drawing for 3D Walkthrough, *11th Annual Symposium on User Interface Software and Technology, ACM UIST'98*, San Fransisco, November 1-4, 1998, pp.173-174.

13. Loomis, J. M., Da Silva, J.A., Fujita, N., & Fukusima, S. S. (1992) Visual Space Perception and Visually Directed Action. *Journal of Experimental Psychology: Human Perception and Performance*, 18, pp. 906-921.

14. Mackinlay, Jock D., Card, Stuart K., and Robertson, George G. A Semantic Analysis of the Design Space of Input Devices. Human-Computer Interaction, 5, 2-3, 1990, pp. 145-190.

15. Mackinlay, Jock D., Card, Stuart K., and Robertson, George G. Rapid Controlled Movement Through a Virtual 3D Workspace. *Conference Proceedings on Computer Graphics*, 1990, pp. 171 – 176.

16. MechWarrior II: The Titanium Trilogy. See http://www4.activision.com/games/mechseries/ for more information.

17. Pausch, Randy, Burnette, Tommy, Brockway, Dan, and Weiblen, Michael E. Navigation and Locomotion in Virtual Worlds via Flight into Hand-held Miniatures. *Proceedings of the 22nd annual ACM Conference on Computer Graphics* , 1995, pp. 399.

18. Peterson, B., Wells, M., Furness, T., and Hunt, E. The Effects of the Interface on Navigation in Virtual Environments. *In Proceedings of Human Factors and Ergonomics Society 1998 Annual Meeting*, 1998, pp. 1496-1505.

19. Pierce, Jeffrey S., Conway, Matthew, van Dantzich, Maarten, Robertson, George. Toolspaces and Glances: Storing, Accessing, and Retrieving Objects in 3D Desktop Applications. In *Proceedings of the 1999 symposium on Interactive 3D Graphics*, 1999, pp. 163-168.

20. Quake II. Visit http://www.idsoftware.com/ for more information.

21. Robertson, George, Czerwinski, Mary, and van Dantzich, Maarten. Immersion in Desktop Virtual Reality. *Proceedings of the 10th annual ACM symposium on User interface software and technology*, 1997, pp. 11 – 19.

22. Strommen, Erik. Children’s Use of Mouse-Based Interfaces to Control Virtual Travel. *Human Factors in Computing Systems*, April 1994, pp. 405-410.

23. van Dam, Andries, Herndon, Kenneth, and Gleicher, Michael. The Challenges of 3D Interaction. *Proceedings of the CHI '94 Conference Companion on Human factors in Computing Systems* , 1994, pp. 469.

24. Ware, C. and Osborne, S. Explorations and Virtual Camera Control in Virtual Three Dimensional Environments. *Computer Graphics*, 24(2), 1990 pp. 175-183.

25. Ware, C. and Slipp, L. Using velocity control to Navigate 3D Graphical Environments: A comparison of Three Interfaces. Proceedings of the Human Factors Society 35th Annual Meeting, pp. 300-304.

26. Zeleznik, Robert, and Forsberg, Andrew. Unicam – 2D Gestural Controls for 3D Environments. In *Proceedings of the 1999 symposium on Interactive 3D Graphics*, 1999, pp. 169-173.