VIRTUAL CONTACT: THE CONTINUUM FROM PURELY VISUAL TO PURELY PHYSICAL

Robert W. Lindeman
Department of Computer Science
The George Washington University
Washington, DC, USA
gogo@gwu.edu

Providing appropriate cues to users when interacting with objects in immersive virtual environments (IVEs) is a difficult task. In addition to individual user differences, environmental factors, and task-specific requirements, the technological complexity of the current state of the art in haptic feedback further increases the difficulty. Though the technology continues to improve, we are still a long way from having haptic feedback that meets the demands of a "general solution" to the problem. This paper focuses on ways of providing effective contact cues in IVEs, starting with purely-visual approaches and moving along a continuum to the use of actual physical objects as high-fidelity interfaces.

INTRODUCTION

To provide high-fidelity experiences in virtual environments, the mantra of system designers when determining the amount and type of feedback to provide to users has typically been: *the more, the better*. In terms of visuals, this means more bits-per-pixel, more pixels per image, more triangles per model, and more textures per scene. In terms of audio, this means more audio channels, more bits-per-channel, higher sampling rates, and 3D spatialization. With regard to interaction, this means "instantaneous" response to user actions (*e.g.*, head movements) and increased "naturalness." For haptic feedback, this means gross characteristics, such as feeling the weight of moveable objects and arresting movement when contacting immovable objects, as well as fine characteristics, such as high-resolution haptic textures, edges and corners of rigid objects, compliance of deformable objects, and friction.

Systems implementing this mantra have met with more or less success, depending on factors such as the task being performed, available computing resources, cost, and target user. By far, the most difficult to address term in this equation has been in providing "adequate-fidelity" in haptic cues. In this paper, we limit our focus to feedback cues used for conveying information about contact the user makes with objects in an immersive virtual environment (IVE). As the adequacy of the cues depends on several factors, we can attempts to provide some qualitative approaches to the effective combination of multimodal cues.

Virtual Contact

We define virtual contact as the user coming into contact with objects in the virtual world (Lindeman *et al.*, 2002). This contact can either be direct, such as grabbing a doorknob or backing into a wall, or indirect, such as bumping a hand-held stylus into a virtual object or manipulating virtual objects from afar. Virtual contact research addresses the problem of what feedback to provide in these situations. Since current technology limits the fidelity we can experience in virtual worlds, it is of interest to us to help define the subset of cues for these environments that will maintain or improve human performance, and to find ways of effectively presenting the stimuli, thereby allowing the higher-level cognitive systems to construct a reality that, though of a lower fidelity, still produces an equivalent experience.

Types of Contact

There are number of ways that we interact physically with objects in the real world. Intuitively, it seems we would like to provide cues in the virtual environment that vary in similar ways. We identify at least two types of contact: impulse and continuous. Impulse contact refers to ballistic interaction, such as knocking on a door or walking down a hallway.

Continuous contact is a more-common occurrence, and refers to situations where we maintain contact over a period of time. This type can be further broken down into the sub-classes of sliding and pushing or pulling.

Sliding. Sliding one object over another, such as our hand along the edge of a table, involves computing forces for satisfying penetration constraints, as well as for resultant friction, which influence our perception of the contact. Surface texture will also have an effect on the tactile experience.

Pushing/Pulling. Pushing or pulling on objects results in changes in forces that are a product of the exerted force, the weight of the object, whether the object is moveable or immoveable, deformability, and/or object movement constraints. In addition, these properties might change over the lifetime of the contact.

Object Properties

An alternative approach is to look at contact from the point of view of the material properties of the objects involved in the contact. The user may be exploring the environment to learn more about the objects contained therein. Surface compliance, object weight, *etc.*, tell us something about the makeup of objects, as do tactile properties such as surface texture and friction.

There are other interesting object properties that can come into play when providing contact cues. When interacting with liquids, viscosity should be taken into account to allow the user to differentiate between liquids. Thermal properties also provide us with information that can be vital to successfully performing tasks.

When attempting to provide adequate cues for contact, researchers have employed techniques which include stimuli to various senses. In the next section, we motivate the need to study the influence of each modality in isolation, as well as in concert.

CHOOSING THE RIGHT MAPPINGS

As humans, we interact with the world using cues from multiple sensory channels, all coordinated to help us make sense of things. The limited multimodal feedback in current IVE systems, however, hinders users from fully understanding the nature of contacts between the user and objects in these environments. Given this situation, some researchers have devised various approaches to improving cues delivered to the user. Many approaches are effective, but task specific, while others, though more general, prove less effective than the former when performing specific tasks.

We can summarize a number of approaches to providing cues for virtual contact. Table 1 gives an overview of some of these techniques. This is by no means an exhaustive list, and we will continue to develop this table as part of this research, in addition to exploring its use.

Table 1: Mapping of Cues to Information

Cue Technique	Modality	Mapped to…
Color change	Visual	Location, depth of penetration
Vector glyphs	Visual	Force and direction of contact
Texture distortion	Visual	Location, depth of penetration
Shape distortion	Visual	Location, depth of penetration
Contact illumination	Visual	Location of collision
Pitch change	Auditory	Depth of penetration
Amplitude change	Auditory	Force of collision
Spatialization	Auditory	Location of collision
Force reflection	Haptic	Weight, surface compliance
Vibration intensity	Haptic	Depth of penetration
Vibration location	Haptic	Location of contact

Visual Approaches

Many systems provide only visual cues when contact is made with virtual objects. The most common of these is the use of a binary color change (*i.e.,* on/off) upon contact. Desktop windowing systems commonly use visual cues, such

as a visual ridge, to show which interface element the cursor is "contacting." This helps in immersive environments as well, in particular as an aid for object selection and manipulation. This method can be further refined by localizing the visual change to the points of contact by altering the color at local vertices, manipulating textures, *etc.* For deformable objects, the amount of deformation can be characterized by smooth variations in the color from the most-deformed to the least-deformed areas of the object.

Another visual approach has to do with imposing object penetration constraints (Lindeman *et al.*, 1999). When interacting with virtual objects, there is a tension between the need to maintain the registration of a tracked object with its real-world counterpart, and the desire to satisfy visual penetration constraints. For instance, if the user's hand is being tracked, and the hand is moved through a virtual wall, should the visual representation of the hand remain at the wall boundary, thereby enforcing the reality that hands do not usually pass through walls, but breaking the registration of the real and virtual hands, or should the visual hand remain registered with the real hand, enforcing the proprioceptive sense, but not the intuition about hands passing through walls? We have shown previously that the answer is relative. Based on the depth of penetration, the former is preferred at shallow penetration depths, while the former is preferred for deep penetration (Lindeman, 1999).

A user touching an object produces sensation in the fingertips. We could use very local illumination to visually illuminate the area around a contact points when a user touches an object. This could be done in a similar fashion to the phosphor that is produced by algae that can be seen at night in the ocean. Phosphor particles could be illuminated, or, alternatively, a local flood light could be used.

Another approach that might prove effective employs glyphs. For example, displaying a force vector at the point of contact might give the user a better understanding of the nature of the contact. It is important to study as many different types of visualization, sonification (Foner, 1999), *etc.* techniques as we can, in order to have some confidence that we will identify usable ones.

A point that needs to be made clear is that most of the feedback we are describing here is delivered exclusively to the person making contact. Because a person making contact with a wall in the real world is the only one aware of the contact, we want to retain this characteristic of contacts. If, however, a user applies adequate force to an object such that a sound would be produced (*e.g.*, kicking in a door), this sound should be propagated to other occupants. We view this as an interesting issue, requiring further study. To our knowledge, little work has been done to determine which cues to "broadcast" to others, and which to keep personal.

The Use of Sound

The most common addition to visuals for providing contact cues in virtual environments has been to use sound. These approaches range from an arbitrary tone being triggered upon contact, as is done in many virtual environments, to a canned sound file being played, such as the "grunt" sound that

is triggered in some video games that is supposed to simulate a person bumping into a wall. Sometimes the audio is preprocessed before it is output to take into account environmental effects of the surrounding virtual space, such as the material properties of the walls, the velocity of the contact, location of the contact, *etc.* (Takala & Hahn, 1992). The footfalls of a video game character as she runs through a tomb, for instance, sound different from when she runs down a marble-floored hallway. This type of cue is very effective at supporting realism.

In addition to impulse-type contact sounds, continuous contact sounds, such as sliding a mug on the surface of a table, can also provide significant contact cues. For instance, a base sound could be associated with each object in the scene, and then when one object contacts another, the base-sounds of the two objects, their friction properties, and material properties of the surrounding environment could all be taken into account to produce a resulting sound that is closer to what might be expected in a real environment.

Finally, sound can also be used to convey information about the deformation of an object. For example, the pitch of a sound could be varied based on the amount of deformation that has been applied to an object, or a stretching sound could be played during the deformation, and then stopped once the action stops.

Haptic Cues

The successful addition of haptic cues to interaction in IVEs has proven to be difficult for the general case. Early systems used mechanical linkages under computer control to provide cues to the hand (Brooks *et al.*, 1990; Iwata, 1990). Though these initial systems occupied large spaces, a reduction in size (Massie & Salisbury, 1994), has not yet sufficiently reduced the cumber inherent in the mechanical approach to allow the freedom of movement necessary for applications requiring the participant to move around freely, such as those typically deployed in cave-like systems.

Where these systems have been successful is in single-point, tool-based applications such as surgical simulation, and shape modeling (Taylor *et al.*, 1993). Scaling these systems up to multiple points of contact, such as the CyberGrasp system from Immersion Corp., has been less successful, because of the added cumber, cost, and computational requirements.

For conveying tactile information, actuated arrays of pins (Bensmaïa & Hollins, 2000) or electro-cutaneous stimulators (Kajimoto *et al.*, 2001) have been used successfully in transmitting cues when a computer cursor crosses some interface boundary. These approaches are limited almost exclusively to the fingerpad of the index finger, where mechanoreceptor density and type are optimal for such techniques.

Vibration

An approach to providing contact cues that has proven very cheap and easy to deploy is the use of vibration. Every major game console on the market today provides some form of vibration feedback. Game designers incorporate use of this technology for various effects, such as for road quality in vehicle-based games, being hit by weapons fire in combat games, bumping into walls in first-person shooters, and even to feel the vibration of nearby heavy machinery. When married with the complementary visual cues necessary to convey the current situation, these vibrational cues can be very effective in providing the subtleties necessary to willingly suspend disbelief.

Though vibrotactile systems have typically been limited to providing feedback to the hands through game controllers, some researchers have begun to look at the possibility and effectiveness of providing feedback to other parts of the body. Rupert with the U.S. Navy and van Erp and his group at TNO in the Netherlands have performed extensive work into the use of vibrotactile cues for conveying the down-vector to pilots (Rupert, 2000; van Erp, 2000).

Kume *et al.* (1998) introduced vibrotactile stimulation on the sole of the foot, and developed a slipper-like interface. They put two stimulators on each sole and made use of phantom sensations. They measured the characteristics of the phantom sensation psychophysically, and found that the location, movement, and rotation of objects could be perceived.

Yano *et al.* (1998) developed a suit-type vibrotactile display with 12 stimulators attached to the forehead (1), palms (2), elbows (2), knees (2), thighs (2), abdomen (1), and back (one on the left side and one on the right). They examined the effectiveness of using this vibrotactile display for tasks that required the user to walk around a virtual corridor visually presented in a cave-like display. They showed that presentation of tactile cues was effective for imparting collision stimuli to the user's body when colliding with walls.

In our own work, we have looked at determining the limits of perception of the human back in terms of vibration intensity and location discrimination (Lindeman & Yanagida, 2003) and as a means for directing the user's gaze for predominantly visual search tasks (Lindeman *et al.*, 2003). We are currently incorporating our findings into the development of an upper-body garment for use in IVEs for conveying contact cues for maneuvering tasks during walking-based locomotion.

Passive-Haptic Feedback

In each of the preceding approaches, the cues provided were all under computer control, and as such, the level of fidelity possible with them is limited by available processor resources. An alternative approach to providing haptic cues is through the use of passive-haptic "devices" (Lindeman, 1999). These approaches use haptic and tactile properties inherent in physical objects to convey high-fidelity, special-purpose feedback.

In an application for treating the fear of heights, Hodges *et al.* (1995) incorporated a physical railing and five-inch platform registered with visual artifacts to enhance the sense of presence in a glass elevator environment. Because the target users suffered from vertigo, reaching for, and firmly holding onto, the railing was a common action, so the presence of the physical counterparts had great effect.

Researchers at UNC-Chapel Hill used Styrofoam blocks to quickly re-create physical walls or counters found in virtual spaces, and allowed users to feel the objects as they explored the space (Insko, 2001). In a comparison of IVEs with and without passive-haptic cues, they found that a faster completion time and a significantly lower number of collisions in navigation of a real-world environment post-test were registered with users who practiced in an IVE with passive haptics versus those who practiced in an IVE without passive haptics. However, they found no difference in estimating spatial properties such as object height or sketching a map of the space. This seems to suggest that passive haptics can be effective for users to learn general navigation skills.

In a very-compelling environment, Brooks and his team provide a physical platform registered in space with a ledge in the virtual environment around a 20-foot pit to enhance the sensation the immersed user experiences (Insko, 2001). Even though the actual ledge is only two inches off of the floor, the simple ability to feel the ledge with the toe of the shoe reinforces the visual aspects of the scene.

In our own work (Lindeman, 1999), we compared the use of visual+audio feedback cues with various additional cues, including the imposition of interpenetration constraints (called *clamping*), and the use of passive haptics, in support of typical interface widget manipulation tasks, such as drag-and-drop and slider manipulation, within IVEs. While the use of clamping significantly improved overall task performance in terms of speed and accuracy compared to letting objects simply pass through others, the presence of passive haptics proved to be superior still for manipulations requiring high precision. These findings have been verified by similar work done in rear-projection environments (Szalavári & Gervautz, 1997).

Feedback Modality Interaction

It should be noted that the interaction between cues fed to multiple channels, be it enhancing or hindering, is a very important issue, but one that is illusive to researchers hoping to control the test environment. Indeed, multiple types of intrasensory feedback can also have a conflicting or supportive affect on user perception. Most research has looked at individual modalities in isolation, as this approach is more tractable for user-study design. Confidence in generalizing results from studies of this kind to specific application areas is sometimes low, because of other feedback modalities that may be present in the system.

When designing user interface studies, researchers are faced with a dilemma in terms of certain interface decisions: we must try to hold constant all aspects of the interfaces that are not being tested. Unfortunately, this means that some of our decisions may skew the results in favor of some interfaces over others. Alternatively, each interface can be designed to approach the optimal interface for the given independent variables. A threat to this method is that we may now be comparing apples to oranges; in other words, it is difficult to make authoritative statements about the influence of the dependent variables, because other factors may have unduly influenced performance measures.

The former approach can be viewed as being more theoretical, while the latter approach as more applied. We encourage researchers in this field to design experiments using both the single modality variation approach, as well as the mixed modality approach.

THE CONTRIBUTION OF INDIVIDUAL CUE QUALITY TO THE OVERALL EXPERIENCE

One of the topics we are looking at in our current research is the contribution that the quality of each individual modality has on the overall quality of the experience. For instance, the visual quality of images and the framerate that can be produced using commodity video cards today is arguably acceptable for most IVE experiences. Audio technology, though far ahead of where it was ten years ago, has not kept pace with video quality in terms of the cost/performance ratio. Haptic feedback systems are even less developed in terms of the cost required to provide "adequate" quality.

Our current working hypothesis for delivering adequate cues is that the contribution of each individual modality is *additive* in relation to the quality of the overall experience, once a given minimum quality threshold for the modality has been reached. As an example, given a stereo visual scene running at 30 frames per second per eye, with negligible head-movement lag, providing rudimentary collision audio cues, such as a non-spatialized impulse tone, will significantly enhance the user's perception of the contact. Looking at this example graphically, we get something like Figure 1.
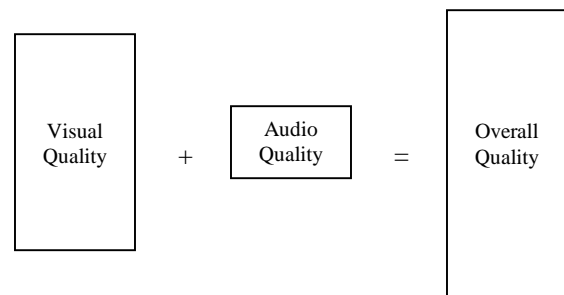

Figure 1: Additive Relationship of Modality Contribution

This working model provides some interesting possibilities for experimentation. We hypothesize that we can provide similar overall quality by varying the contribution of each modality. Because quality is typically limited by current technology, *e.g.*, GPU quality, audio subsystem, haptic rendering rates, it would be convenient to have some way to assess the quality of the overall experience possible for a given set of available resources.

In the haptic domain specifically, this would allow us to assess the contribution of such "cheap tricks" as the use of Styrofoam blocks for walls, or vibrotactile devices for whole-body contact, in place of more-cumbersome mechanical systems. Though this work is still in its early stages, we see great potential in quantifying the user experience.

CONCLUSION

In this paper, we have attempted to outline some of the issues that need to be addressed when designing contact cues for users in IVEs as they interact with virtual objects. Within each modality, we have identified several possible ways that information about contact can be conveyed. In addition, we can combine feedback from multiple modalities, and here we attempt to provide a way of defining the overall quality of the experience commensurate with available resources.

References

Bensmaïa S., Hollins M. Complex tactile waveform discrimination. *J. of the Acoustical Soc. of America*, 108(3), 2000, pp. 1236-1245.

Brooks, F.P., Ouh-Young, M., Batter, J.J., Kilpatrick, P.J. Project GROPE - Haptic displays for scientific visualization. *Proc. of ACM SIGGRAPH '90*, 1990, pp. 177-185.

Erp, J.B.F. van. Tactile information presentation: Navigation in virtual environments. *Proc. of the First International Workshop on Haptic Human-Computer Interaction*. Brewster, S., and Murray-Smith, R. (Eds.), Glasgow, UK, August/September 2000, pp. 165-173.

Foner, L.N. Artificial synesthesia via sonification: a wearable augmented sensory system. *Mobile Networks and Applications*, 4(1), 1999, pp. 75-81.

Hodges, L.F., Rothbaum, B.O., Kooper, R., Opdyke, D., Meyer, T., North, M., de Graaff, J.J., Williford, J. Virtual environments for treating the fear of heights. *IEEE Computer*, 28(7), 1995, pp. 27-34.

Insko, B.E. Passive haptics significantly enhances virtual environments. Doctoral dissertation, Dept. of Computer Science, University of North Carolina at Chapel Hill, 2001.

Iwata, H. Artificial reality with force-feedback: development of desktop virtual space with compact master manipulator. *Proc. of ACM SIGGRAPH '90,* 1990, pp. 165-170.

Kajimoto, H., Kawakami, N., Maeda, T., Tachi, S. Electrocutaneous display as an interface to a virtual tactile world. *Proc. of IEEE Virtual Reality 2001*, 2001, pp. 289-290.

Kume, Y., Shirai, A., Tsuda, M., Hatada, T. Information transmission through soles by vibrotactile stimulation. *Trans. of the Virtual Reality Society of Japan,*. 3(3), 1998, pp. 83-88.

Lindeman, R.W., Bimanual interaction, passive-haptic feedback, 3D widget representation, and simulated surface constraints for interaction in immersive virtual environments. Doctoral dissertation, Dept. of EE & CS, The George Washington University, Washington, DC, 20052, May 16, 1999.

Lindeman, R.W., Sibert, J.L., Hahn, J.K. Towards Usable VR: An Empirical Study of User Interfaces for Immersive Virtual Environments. *Proc. of ACM CHI '99*, 1999, pp. 64-71.

Lindeman, R.W., Templeman, J.N., Sibert, J.L., Cutler, J.R. Handling of Virtual Contact in Immersive Virtual Environments: Beyond Visuals. *Virtual Reality*, 6(3), 2002, pp. 130-139.

Lindeman, R.W., Yanagida, Y. Empirical Studies for Effective Near-Field Haptics in Virtual Environments. *Proc. of IEEEVirtual Reality 2003*, 2003, pp. 287-288.

Lindeman, R.W., Yanagida, Y., Sibert, J.L., Lavine, R. Effective Vibrotactile Cueing in a Visual Search Task. To appear in *Proc. of Interact 2003*.

Massie, T.H., Salisbury, J.K. The PHANToM Haptic Interface: A Device for Probing Virtual Objects. *Proc. of the ASME WAM, Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, DSC–Volume 55-1, 1994, pp. 295-300.

Rupert, A. An instrumentation solution for reducing spatial disorientation mishaps. *IEEE Eng. in Med. and Bio. 2000*, pp. 71-80.

Szalavári, Z., Gervautz, M. The personal interaction panel - a two-handed interface for augmented reality. *Proc. of Eurographics'97*, 16(3), 1997.

Takala T., Hahn J. Sound Rendering. *Proc. of ACM SIGGRAPH '92*, 1992, pp. 211-220.

Taylor, R.M., Robinett, W., Chi, V.L., Brooks, F.P., Wright, W.V., Williams, R.S., Snyder, E.J. The Nanomanipulator: A Virtual-Reality Interface for a Scanning Tunneling Microscope. *Proc. of SIGGRAPH '93*, 1993.

Yano, H., Ogi, T., Hirose, M. Development of Haptic Suit for Whole Human Body Using Vibrators. *Trans. of the Virtual Reality Society of Japan*, Vol. 3, No. 3, 1998, pp. 141-148.