Multimodal Virtual Reality: Input–Output Devices, System Integration, and Human Factors

Grigore Burdea

Rutgers—The State University of New Jersey

Paul Richard
Philippe Coiffet

University of Paris VI

Virtual reality (VR) involves multimodal interactions with computer-simulated worlds through visual, auditory, and haptic feedback. This article reviews the state of the art in special-purpose input–output devices, such as trackers, sensing gloves, 3-D audio cards, stereo displays, and haptic feedback masters. The integration of these devices in local and network-distributed VR simulation systems is subsequently discussed. Finally, we present human-factor studies that quantify the benefits of several feedback modalities on simulation realism and sensorial immersion. Specifically, we consider tracking and dextrous manipulation task performance in terms of error rates and learning times when graphics, audio, and haptic feedback are provided.

INTRODUCTION

Virtual reality (VR), also called "virtual environments" (Sheridan, 1992), "cyberspace" (Elmer-Dewitt, 1993), "veridical environments" (Codella, Jalili, Koved, &

Research reported here has been supported in part by grants from the Center for Computer Aids for Industrial Productivity (CAIP) at Rutgers—The State University of New Jersey and by the Centre de Robotique Integree d'Ile de France. We gratefully acknowledge the participation of many volunteers who took part in human-factors experiments.

Correspondence and requests for reprints should be sent to Grigore Burdea, Rutgers University—CAIP Center, P.O. Box 1390, Piscataway, NJ 08855–1390.

E-mail: burdea@caip.rutgers.edu

Lewis, 1993), or "artificial reality" (Krueger, 1991), represents a high-end graphics user interface that immerses the user in a simulated world. In the case of VR or veridical environments, the simulated world aims at replicating the known physical reality. Conversely, for artificial reality and virtual environments, the simulation may depart from known reality, allowing for experiences that cannot be reproduced in the real world.

Key to the user's VR experience are *multimodal real-time interactions* involving all human senses, from vision (Robinett & Rolland, 1992) and hearing (Wenzel, 1992), to tactile (Marcus, 1993), smell (Keller, Kouzes, Kangas, & Hashem, 1995), and even taste (Bardot et al., 1992). This rich sensorial interaction (sometimes even sensorial overload), coupled with real-time simulation responses produces a compelling and captivating feeling of immersion. Sheridan (1992) referred to this feeling: "In some ideal sense, and presumably with sufficiently good technology, a person would not be able to distinguish between actual presence, telepresence, and virtual presence" (p. 274).

In order to produce simulations that respond in real time to user's input, it is necessary to have powerful computing platforms. This requires fast processors to handle input–output communications, task scheduling, and physical modeling. Large graphics loads involving tens of thousands of polygons that are Gouraud shaded (Gouraud, 1971) or textured (Foley, Van Dam, Feiner, & Hughes, 1990) require additional graphics accelerator cards. Such cards range from the "Reality Engine" (Silicon Graphics Inc., personal communication, 1992) for SGI workstations, to the "Fire Board" for the PC (Spea Video Seven, 1992). Producing the same graphics in stereo doubles the graphics load and normally requires two such graphics accelerators (one for each eye). The computer together with its graphics accelerators and (sometimes) math coprocessors form the "VR engine." It receives input from the user through input–output devices and uses on-board databases and software libraries to render and display the simulated world. The components of a VR system are illustrated in Figure 1 (Burdea & Coiffet, 1994).

Looking at the figure it becomes clear that all user-computer interactions are mediated by the input–output devices. The utility of a VR system for application development and user training is ultimately influenced not only by the computing platform and available software tools, but also by the quality of these specialized VR input–output devices. By "quality," we mean ease of installation and use, preservation of user safety and freedom of motion, and minimization of user fatigue. Additional input–output tool design requirements are matching the human sensorial characteristics, high measurement and feedback signal bandwidth, low sampling and transmission latencies, and low signal noise. These qualities result in an interface that is "transparent" and a simulation that is responsive and feels "natural."

Input–output devices and the specific sensorial modalities they are mediating are the topic of this article. Section 2 describes available input–output device technology from trackers and sensing gloves to the latest haptic (force and tactile) interfaces. Section 3 describes VR systems that integrate both nonhaptic and haptic sensorial modalities in a single-user distributed simulation. Human-factor studies that quantify the influence of various feedback modalities on task error

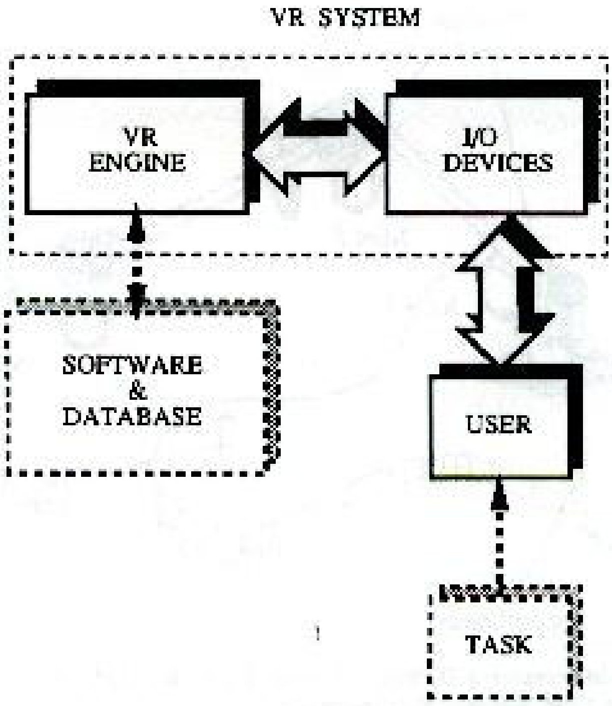

Figure 1. VR System block diagram (Burdea & Coiffet, 1994). Reprinted by permission. Copyright 1994, Hermes Publishing Co.

rates and learning times are described in Section 4. Concluding remarks are given in Section 5.

INPUT–OUTPUT DEVICES

Position Trackers

The VR simulation loop generally starts by the sampling of user head and hand positions that is done at rates of 20–120 measurements/s. These measurements are data sets of six numbers representing three-dimensional (3-D) translations and orientations and are measured by dedicated devices called "trackers." If the hand or head position is measured with respect to a fixed system of coordinates, then the tracker is "absolute." Otherwise, only an incremental motion is detected and the tracker is "relative." Examples of absolute trackers are the Polhemus Fastrack (Krieg, 1993), the "Flock of birds" (Scully, 1993), or the Logitech ultrasonic tracker (Sowizral & Barnes, 1993). Relative trackers are trackballs or joysticks that measure forces (or displacements) applied by the user on a compliant element (Bowman, 1993; Hirzinger, 1987). Present absolute trackers

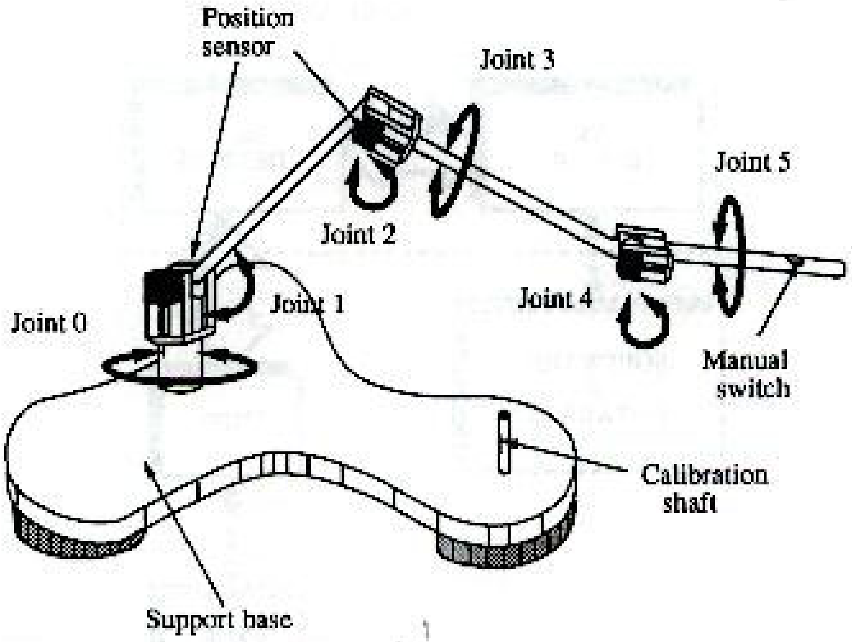

Figure 2. The Immersion 3-D Probe (Burdea & Coiffet, 1994). Reprinted by permission. Copyright 1994, Hermes Publishing Co.

have a small measurement range, sampling rates that depend on the number of probes used and accuracies that are affected by magnetic fields or ultrasonic noise. Relative trackers do not have these problems, but the user's freedom of motion is sacrificed by the need to keep in contact with the desk-based trackball or joystick.

A newer device that is a hybrid between an absolute and relative tracker is the Immersion "3-D Probe" (Immersion Co., 1993) illustrated in Figure 2 (Burdea & Coiffet, 1994). It consists of a small mechanical arm that integrates position sensors at its joints. By direct kinematics calculation, it is possible to measure the absolute position of the user's hand that holds the probe versus the arm base. The overall simulation volume is thus much larger than for a trackball. For even larger volumes (required for example during fly-by simulations), the absolute position can be interpreted as a relative increment and the probe becomes a relative tracker. Because no magnetic or ultrasonic signals are used, the 3-D Probe is immune to magnetic or ultrasonic interference, while its sampling rate is an impressive 200 positions/s.

Sensing Gloves

Applications that involve highly dextrous manipulations of the virtual environment require devices that measure finger (rather than wrist) positions. Such devices are sensing gloves like the "5th Glove" (Roehl, 1995a) and the "CyberGlove" (Kramer, Lindener, & George, 1991) or exoskeletons such as the "Dextrous Hand Master" (Marcus, Lawrence, & Churchill, 1991). These are devices worn on the user's hand that measure finger joints at rates of about 30 data sets/s. These joint values are then used in combination with the wrist position to determine hand gestures. One problem with sensing gloves is their more complicated calibration, required by user-specific hand characteristics. Another drawback is higher hardware prices. Additionally, few of the commercially available gloves have haptic (force or tactile) feedback that is a drawback for realistic simulations.

The aforementioned trackers and gloves input commands into the simulation environment. These commands are interpreted by the computer in conjunction with geometrical, physical, and behavioral models of various virtual objects. What results is a change in the state of the virtual world that needs to be presented (or fed back) to the user. All VR systems use visual (or graphical) feedback, to which some add audio and haptic feedback.

Visual Feedback

Depending on the type of visual feedback, the VR system may be "fully immersive" or "partially immersive." Fully immersive simulations use "head-mounted displays" (HMDs) which are worn on the user's head. These HMDs have two screens each displaying a separate image of the virtual world. The two images are integrated in the brain to produce a stereo image. Depending on the type of display used HMDs are liquid-crystal-display (LCD) based or CRT based, as illustrated in Figure 3 (Burdea & Coiffet, 1994). LCD-based HMDs have the advantage of lower prices (compared with CRT-based models), compactness, and lightness. Modern systems weigh as little as 0.23 kg for the HMSI-

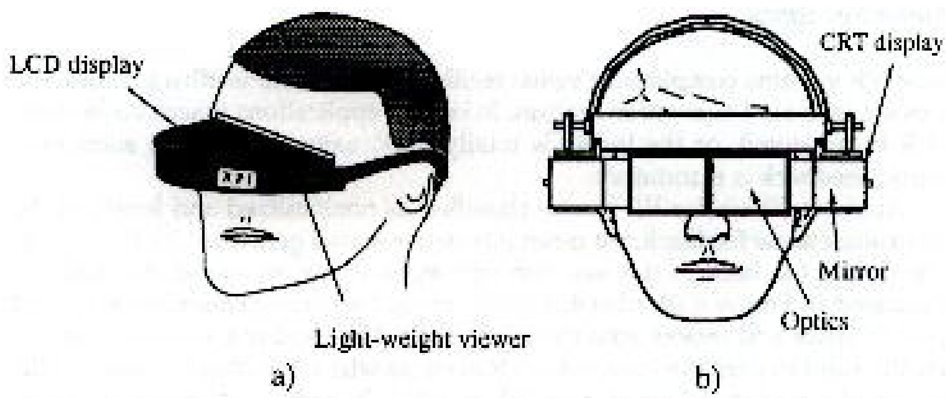

Figure 3. HMDs: (a) LCD based and (b) CRT based (adapted from Burdea & Coiffet, 1994). Reprinted by permission. Copyright 1994, Hermes Publishing Co.

1000 (RPI Advanced Technical Group, 1993), or 8 oz for the Virtual i/o "i-glasses!" (Roehl, 1995b). This represents a clear improvement over earlier models that weighed 2 kg. The drawbacks of compact LCD-based HMDs are low image resolution (about seven times lower than for a graphics workstation) and restricted field of view (30° horizontal field of view vs. 180° human field of view).

When image resolution is critical it is necessary to use high-end CRT-based HMDs. These devices use two miniature CRTs placed laterally to the user's head. The image is reflected by 45° mirrors and then viewed by the user through special optics. Such an HMD is the n-Vision Inc. "Datavision 9c" with a resolution of 1280 × 960 and a 50° diagonal field of view. The drawback with CRT-based HMDs are higher prices compared with the LCD-based models. There are additional safety concerns because high voltages and strong magnetic fields are placed close to the head.

Full immersion may not be required by certain applications (such as CAD modeling or laparoscopic surgery). In this case, it is sufficient to use partly immersive systems based on stereo workstations or projection screens and special glasses (Faris, 1992). Stereo workstations double the hardware refresh rate in order to time sequence left and right-eye images at 120 Hz. The user wears "active" glasses such as the "CrystalEyes" (Akka, 1992) which sequentially block the view of each eye. An infrared controller is used to synchronize the stereo monitor images with the shutters incorporated in the glasses. These systems are light and comfortable and have higher resolution and field of view versus LCD-based HMDs. Projection screens use spatially multiplexed images viewed through a polarizing film and polarizing glasses. These glasses are passive (with no electronics) and have a low cost, allowing for a large number of users to participate in a given simulation. Unfortunately, the polarizing film absorbs some of the reflected light, such that images look dimmer than seen in direct viewing on a monitor.

Audio Feedback

Most VR systems complement visual feedback with audio feedback in order to provide increased simulation realism. In certain applications where visual feedback is corrupted, or the image is totally dark, using audio (and sometimes haptic) feedback is mandatory.

Audio feedback for VR can be classified as nonlocalized and localized. By *nonlocalized audio* feedback, we mean interactive sound generated by the simulation that is fed back to the user through mono or stereo sound channels. As illustrated in Figure 4 (Burdea & Coiffet, 1994), the virtual sound source is not fixed in space and rotates with the user's head. Nonlocalized audio feedback is usually used in partially immersive VR systems where the object producing the perceived sound is always in view (Akka, 1992). In such applications, the small discrepancy between the visual and audio source locations is not critical. Nonlocalized audio is also inexpensive and thus used in low-end VR systems.

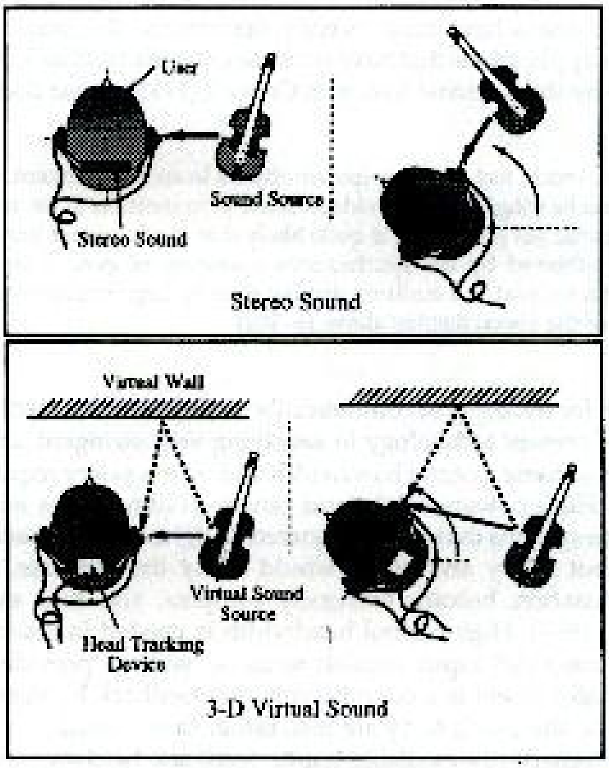

Figure 4. Stereo sound versus 3-D virtual sound (Burdea & Coiffet, 1994). Reprinted by permission. Copyright 1994, Hermes Publishing Co.

In fully immersive VR simulations, part (or all) of the object of interest may not be in the user's field of view. In such instances, it is necessary to provide a sound queue as to where the object is (e.g., a ball bouncing on the floor). In order to produce 3-D audio localization, it is necessary to measure the orientation of the user's head (through a tracker) and combine that with the location of the virtual sound source. The interactive sound is processed through "head-related transfer functions" (Wenzel, 1992), filtered, and convolved such that the perceived location of the virtual sound source is fixed. This processing is done by specialized hardware called "convolvotron," "beachtron," and "acoustetron" which can localize up to 16 audio sources.

Haptic Feedback

Haptic feedback groups the modalities of tactile and force feedback corresponding to forces generated during virtual object interaction. Compared to visual and audio feedback, haptic feedback is a much newer sensorial modality for VR simulations. VR users have only recently understood the major limitations of most of today's applications that have no force or tactile feedback. A recent report commissioned by the National Research Council (1995) stated that:

> Being able to touch, feel, and manipulate objects in an environment, in addition to seeing (and hearing) them, provides a sense of immersion in the environment that is otherwise not possible. It is quite likely that much greater immersion in a VE can be achieved by the synchronous operation of even a simple haptic interface with a visual and auditory display, than by large improvements in, say, the fidelity of the visual display alone. (p. 162)

One reason for the delay in commercially available haptic feedback hardware was the lack of present technology in satisfying very stringent actuator power–weight, power–volume, control bandwidth and user's safety requirements (Burdea, in press). High power–weight and power–volume ratios are necessary in order to provide systems that are both powerful, light, and compact. Systems that are powerful but heavy and bulky would easily tire the user, require active gravity compensation, become extremely complex, and thus expensive (Bergamasco et al., 1994). High control bandwidth is needed in order to match the human tactile sensorial input requirements, as well as provide a responsive simulation. Finally, safety is a concern in haptic feedback because the feedback forces applied on the user's body are real, rather than virtual.

Current commercially available haptic feedback hardware provides either tactile or force feedback but not both. An example of *tactile feedback* is the "Touch Master" (Exos Inc., 1993) that uses piezoelectric miniature disks placed at the user's fingertips. These disks are vibrated at fixed frequencies of 210 Hz and amplitudes that are modulated based on the simulated tactile interactions. The system is compact and relatively inexpensive but cannot reproduce the geometry of touch because only one actuator is used per fingertip. In order to convey geometry information at the fingertip, it is necessary to use compact actuator arrays, such as the micropins shown in Figure 5 (Burdea & Coiffet, 1994). These micropins are usually shape-memory metals (SMM) controlled via an excitation current. By pulsing rows sequentially, it is possible to simulate sliding contact with the edge of a virtual object. Pulsing individual pins temporally attempts to convey surface smoothness information. Such SMM arrays have the advantage of lightness and compactness but have low spatial densities and a very low control bandwidth.

Force feedback masters provide the user with the feel of the compliance (or hardness), weight, and inertia of virtual objects that are manipulated in the simulation. The vast majority of commercially available systems provide force feedback only at the user's hand. Depending on the range of motion allowed, these systems can be further classified as nonportable and portable. Nonportable force feedback is produced by joysticks and small robotic arms that limit the hand range of motion. They generally use electrical actuators that have their weight supported by the desk on which the system is attached. This results in a

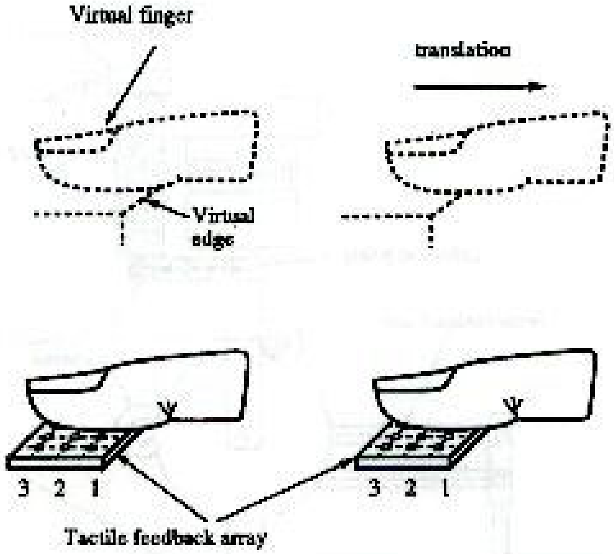

Figure 5. Conveying touch-area geometry information using micropin arrays (Burdea & Coiffet, 1994). Reprinted by permission. Copyright 1994, Hermes Publishing Co.

drastic reduction in the user's freedom of motion. Portable force feedback "masters" are structures worn by the user and thus travel with his or her hand. The resulting freedom of motion and simulation "naturalness" are greatly increased versus nonportable systems.

A recent example of nonportable force feedback hardware is the "PHANToM Master" arm illustrated in Figure 6 (Massie & Salisbury, 1994). It consists of a small 6 DOF robotic arm that is gravitationally counterbalanced and attached to the user's desk. It integrates three DC actuators that provide three translational force feedback degrees of freedom. Force control is done by a PC that reads actuator shaft positions through colocated encoders and an A/D board. The system has the advantage of high control bandwidth, compactness, and easy installation. The drawbacks are limited force capability (approximately 10% of human maximum finger force output) and a limited number of degrees of freedom with force feedback (three).

Certain applications, such as maintenance or assembly training, require the simulation of dextrous manipulation (using the fingers rather than the wrist; Cutkosky & Howe, 1990). In such cases, it is necessary to provide force feedback at independent fingers and allow large simulation volumes. These requirements preclude the use of a joystick or a small robotic arm. One system that allows independent finger force feedback and is fully portable is the Rutgers Master I illustrated in Figure 7 (Burdea & Coiffet, 1994).

The Rutgers Master I (Burdea et al., 1992) is designed to retrofit present "open-loop" sensing gloves by integrating small pneumatic micro-pistons.

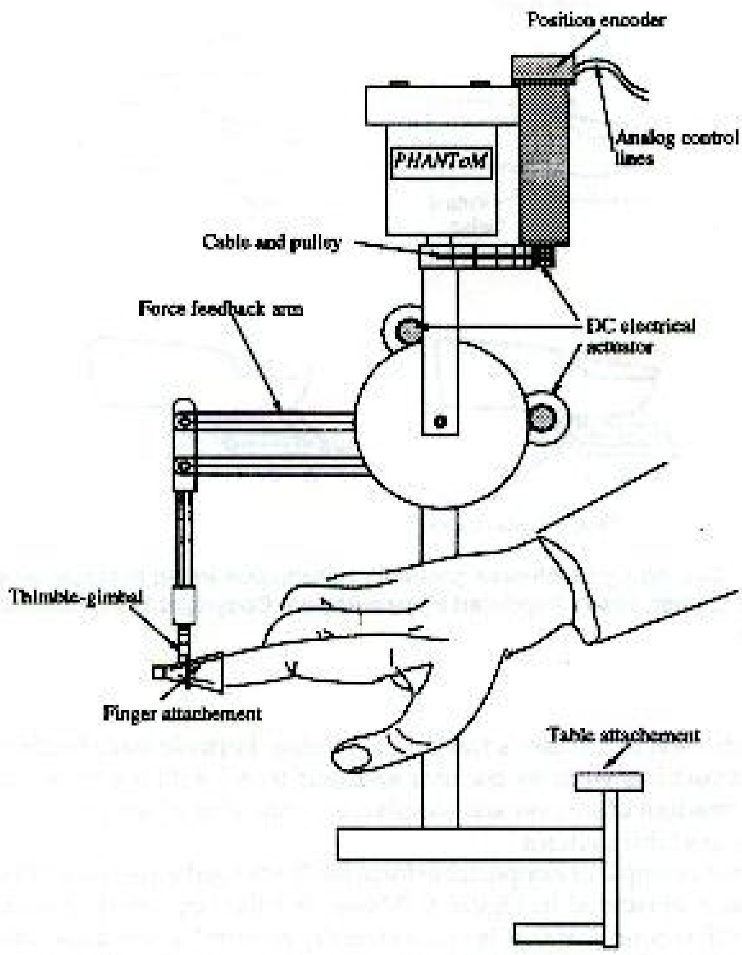

Figure 6. The PHANToM Master (adapted from Massie & Salisbury, 1994). Reprinted by permission. Copyright 1994, ASME.

These actuators are light, have power–weight ratios superior to electrical actuators, and an acceptable control bandwidth. They are also clean and safe (due to air intrinsic compliance and built-in mechanical stops). The actuators are placed on a movable L-shaped platform that allows for variation in user's hand size. Control is done by analog proportional air valves placed in a separate interface box. LED bar graphs are used to visualize the level of forces (actually internal voltages) applied on each finger. This system has been recently redesigned by integrating the fingertip sensing into the force feedback structure (Burdea & Gomez, 1994; Gomez, Burdea, & Langrana, 1994). Thus, the "Rutgers Master II" eliminates the need for a separate sensing glove with advantages in system simplicity and reduced cost.

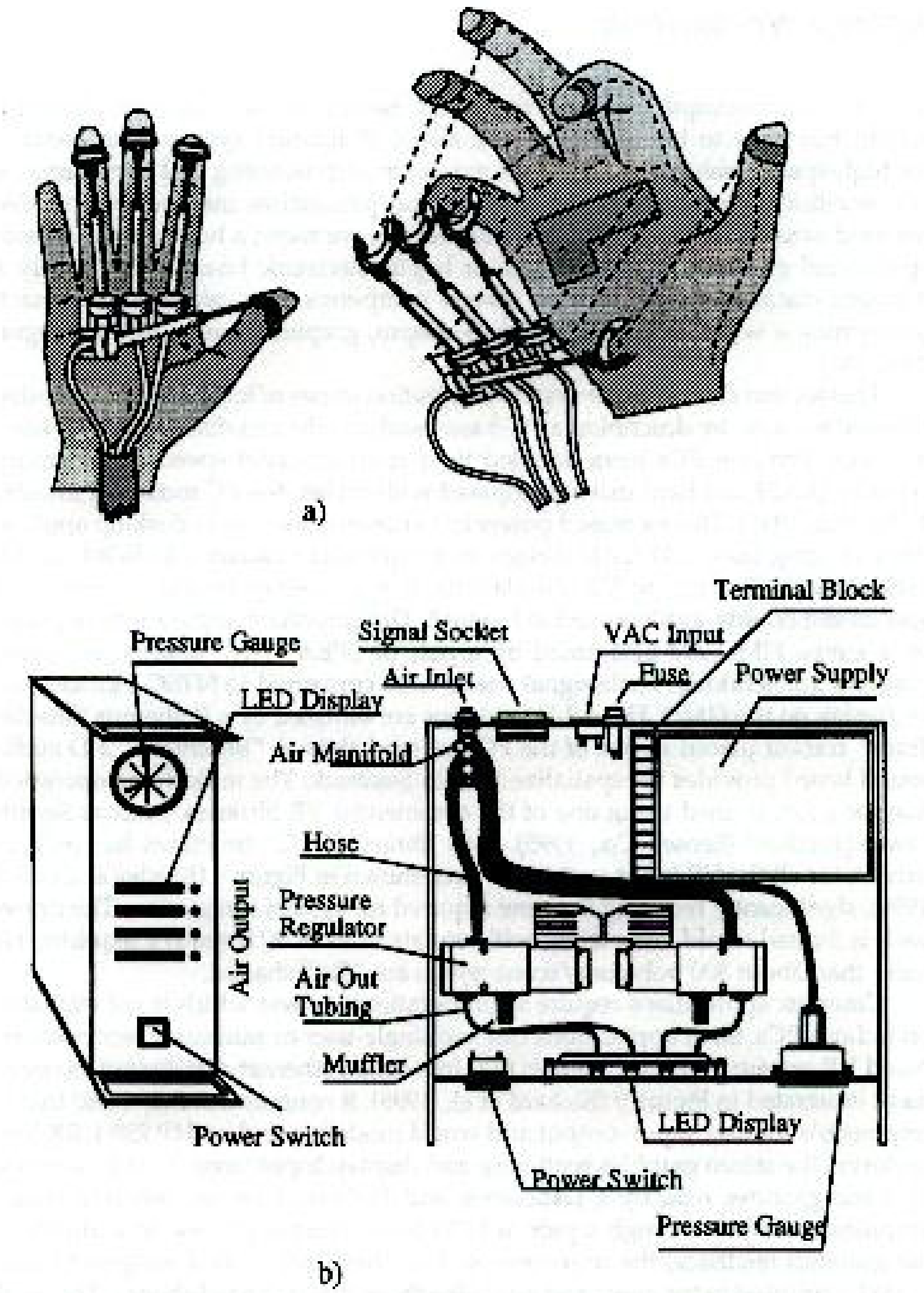

Figure 7. The Rutgers Master I: (a) feedback structure and (b) control interface (Burdea & Coiffet, 1994). Reprinted by permission. Copyright 1994, Hermes Publishing Co.

SYSTEM INTEGRATION

The VR input–output devices described earlier do not operate independently, but have to be integrated in a single simulation system. The demand for high-speed, real-time simulations require multiprocessing and asynchronous computations. Furthermore, the computation parallelism may be local or distributed over the network. By local parallelism, we mean a host computer with specialized graphics, tracker, sound, or haptic electronic boards. Conversely, a network distributed system uses several computers (e.g., workstations), each performing a separate task (user input–output, graphics, audio, voice recognition, etc.).

This section discusses the system integration issues of local and network-distributed systems by describing a PC-based and an ethernet distributed VR environment. Pentium PCs have doubled their computational speed and memory capacity (RAM and hard drive) compared with earlier 486-PC models (Pimentel & Teixeira, 1993). This increased power has allowed their use in desktop applications ranging from 3-D CAD design to multimedia (usually CD-ROM based) presentations. For use in VR simulations, it is necessary to add a number of specialized boards, as illustrated in Figure 8. The graphics computations required by a stereo HMD are performed by a pair of SPEA "Fire" boards (or newer graphics accelerators). Their signal needs to be converted to NTSC format prior to display on the HMD. Head 3-D positions are sampled by a Polhemus "Inside-Track" tracker placed in one of the PC bus card slots. A "Beachtron" 3-D audio sound board provides the spatialized audio feedback. The underlying operation may be programmed using one of the commercial VR libraries, such as Sense8 "WorldToolKit" (Sense8 Co., 1995). This library of "C" functions has built-in drivers for all the VR input–output devices shown in Figure 6 (Burdea & Coiffet, 1994), significantly reducing the time required for system integration. The drawback is limited world complexity, with update rates of 30 frames/s requiring no more than about 300 polygons/scene which are "flat" shaded.

Complex applications require a computational power which is not available on today's PCs. Such applications can use single-user or multiuser workstation-based VR systems. Such a system is the Unix-based ethernet-distributed environment illustrated in Figure 9 (Richard et al., 1996). It consists of a Sun 4-380 that is responsible for user input–output and world modeling and an HP 755-CRX that performs the stereo graphics rendering and display. Input from the user is given by hand gestures read by a DataGlove and Polhemus sensor, whereas stereo graphics is viewed through a pair of LCD-based shutter glasses. In addition to the graphics feedback, the user receives force feedback from a Rutgers Master (RM) I retrofitted to the glove and audio feedback through headphones. The level of feedback forces are visualized by LED bar graphs incorporated in the RM interface box. The underlying software consists of two asynchronous loops, namely the force feedback loop and the graphics loop. Hand positions and graphics scenes are refreshed at approximately 28 frames/s, whereas force feedback bandwidth is approximately 14 Hz.

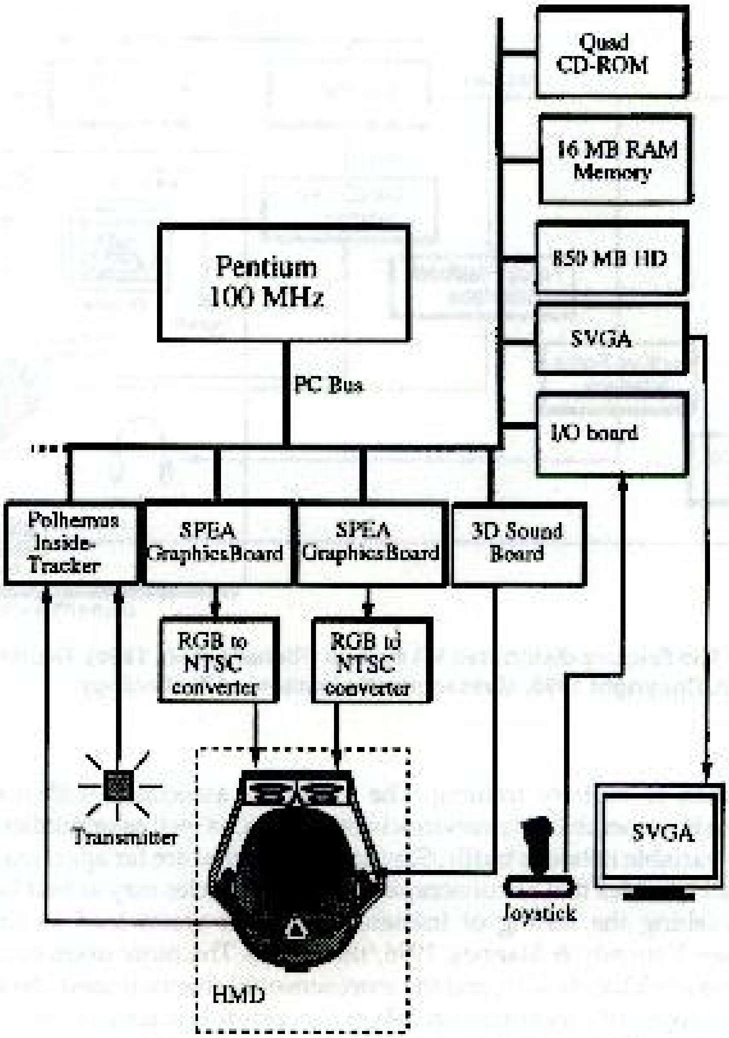

Figure 8. PC-based VR system (adapted from Burdea & Coiffet, 1994). Reprinted by permission. Copyright 1994, Hermes Publishing Co.

The aforementioned distributed simulation system has several advantages versus the PC-based system described previously. First, graphics complexity can now reach 5,000 polygons/scene at 30 frames/s. The second advantage is increased simulation realism through the addition of haptic feedback. Haptics, as well as graphics feedback, can be presented locally or be transmitted remotely over the ethernet. It is thus possible (although not presently implemented in the system described earlier) to have several users interact in the same simulation (see Moshell & Hughes, 1996/this issue). Multiuser capability is an essential requirement for applications ranging from cooperative design, to advanced entertain-

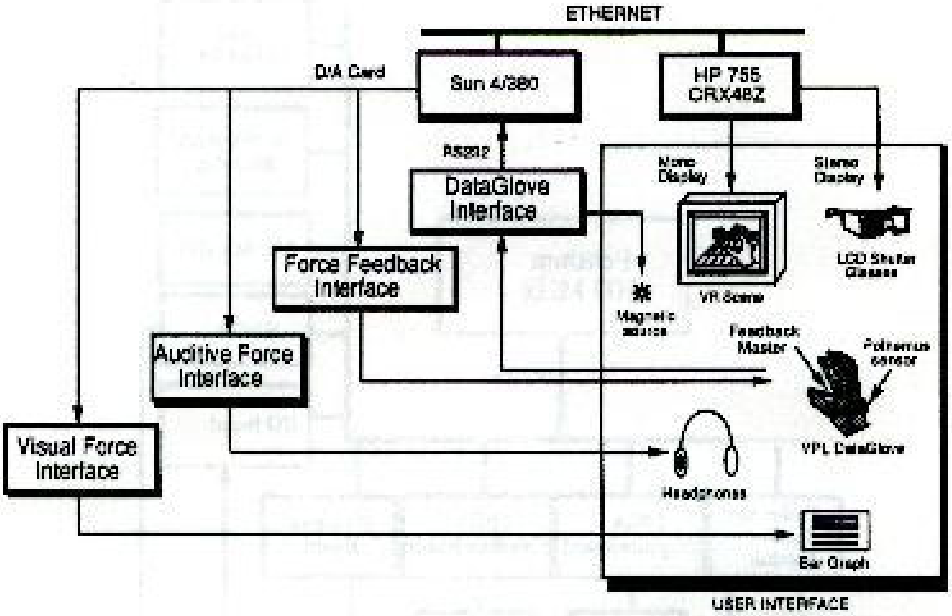

Figure 9. The Rutgers distributed VR system (Richard et al., 1996). Reprinted by permission. Copyright 1996, Massachusetts Institute of Technology.

ment arcades, to military training. The drawback associated with distributed simulations is vulnerability to network breakdowns, as well as unpredictable timing due to variable network traffic. Simulation sites that are far apart may experience overall latencies that are unacceptable. Such latencies may at best be confusing (diminishing the feeling of immersion) and at worst lead to simulation sickness (see Kennedy & Stanney, 1996/this issue). The more users compete for available network bandwidth, and the more sensorial channels used, the more difficult it is to cope with transmission delays. An example is remote force feedback versus tactile feedback transmission. Force feedback may be easily encoded and transmitted over the network with relative ease. However, large tactile arrays require transmission time that grows with the array size, the frequency of excitation, and the number of (intensity) bits for each actuator. One remedy is to have local libraries of tactile patterns and only transmit high-level calls over the network. This approach is taken by Xtensory Inc. for their "tactools" library (Cutt, 1993).

HUMAN-FACTOR STUDIES

The distributed VR simulation shown in Figure 9 was used to perform a series of human-factor studies involving multimodal VR interactions. These studies used an experimental configuration in which users were partly immersed in a virtual

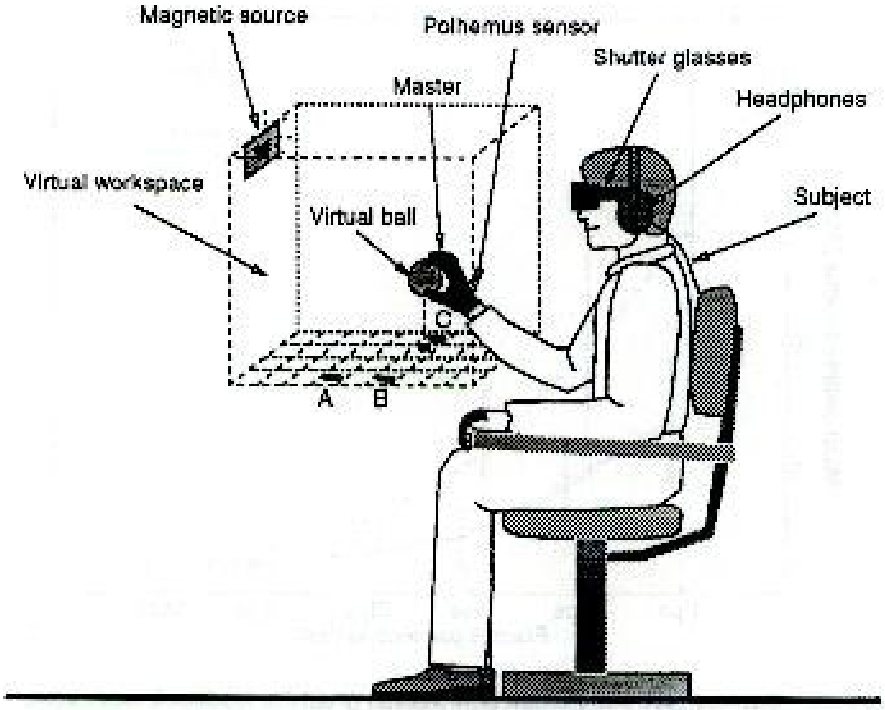

Figure 10. Experimental setup for VR human-factor studies (Richard et al.,1996). Reprinted by permission. Copyright 1996, Massachusetts Institute of Technology.

world consisting of five walls, a deformable ball, and a virtual hand, as illustrated in Figure 10 (Richard et al., 1996).

The first study quantified the influence of graphics modality (stereo vs. mono) and refresh rate (frames/s) on the completion time of a tracking and capture task. A total of 84 participants (42 men and 42 women) were divided into two groups, one viewing a monoscopicly rendered scene, the other a stereo scene. Each group was subdivided into six subgroups (of seven participants each) called group G1 to G6. The graphics refresh rate was group dependent and degraded from 28 frames/s for G1 to only one frame/s for G6. Participants were instructed to track and grab as quickly as possible a ball that appeared at the same location, but with directions varying randomly in a 45° cone. Each session consisted of 10 trials with 15-s rest periods in between. The overall capture time means and standard deviations are plotted in Figure 11. Results showed that there was little degradation when refresh rates were halved from 28 to 14 frames/s, both for monoscopic and stereo graphics feedback. However, grasping time increased logarithmically afterwards, especially for refresh rates lower than 7 frames/s. The standard deviation was also larger, indicating increased subject variability for low refresh rates. Stereo graphics was very beneficial at low refresh rates,

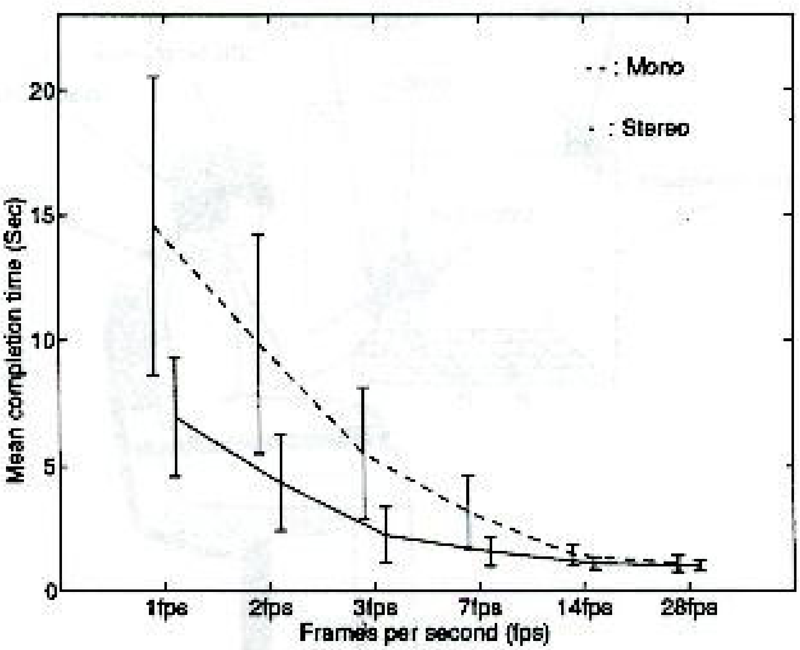

Figure 11. Capture task completion time versus graphics feedback modalities (Richard et al., 1996). Reprinted by permission. Copyright 1996, Massachusetts Institute of Technology.

resulting in a 50% reduction in capture time versus monoscopic scenes. However, at higher refresh rates (above 14 fps) the difference in task completion time was insignificant when using either mono or stereo graphics. Without overgeneralization it seems that there is no need for more expensive stereo HMDs when tasks involve tracking in uncluttered environments. One can use monoscopic displays as long as the graphics workstation can render the virtual scene at reasonable refresh rates.

The second study involved dextrous manipulation of a plastically deformable virtual ball with various feedback modalities. The study used six groups of 14 participants (a total of 42 men and 42 women). The first group (N) had only graphics feedback from the computer screen and the second group (V) had graphics and visual feedback of the contact forces from the LED bar graphs; the third group (A) had graphics and audio feedback, whereas the fourth group (H) had graphic and force feedback using the Rutgers Master. The fifth and sixth groups had redundant force feedback, namely the fifth group had haptic and visual feedback of contact forces (H–V) and the sixth group had haptic and audio feedback of contact forces (Group H–A). Each group was divided in two subgroups, one viewing the scene in mono, the other in stereo. The task consisted of grasping the ball lightly (without deforming it more than 10% of the sphere radius), then moving to a via point "B," and finally releasing it at point

"C." Each session consisted of 10 trials with 15-s rest periods in between. The resulting ball deformation means and standard deviations are plotted in Figure 12. It can be seen that the best results correspond to redundant force feedback (Group H–A) where at the end of the trials (ET) the 10% deformation goal was attained. The worst performance (largest mean and standard deviation) corresponds to group N, which had only graphics feedback. Group H, which had force feedback, did better than Groups N, V, and A but worse than Groups H–V and H–A. This is due to the intrinsic dependence of force feedback on object deformation (required by the Hooke's law $F = k\Delta x$). For small ball deformations Δx the resultant feedback forces were masked by the static friction present in the feedback actuators. The addition of audio feedback consisting of a tone frequency proportional with the contact force was a useful cue of initial contact with the virtual ball. In this way, users minimized the applied forces and improved performance. Additional benefits resulted from task learning, as shown in the reduction in ball deformation from the beginning (BT) to the end of (ET) of the experimental session.

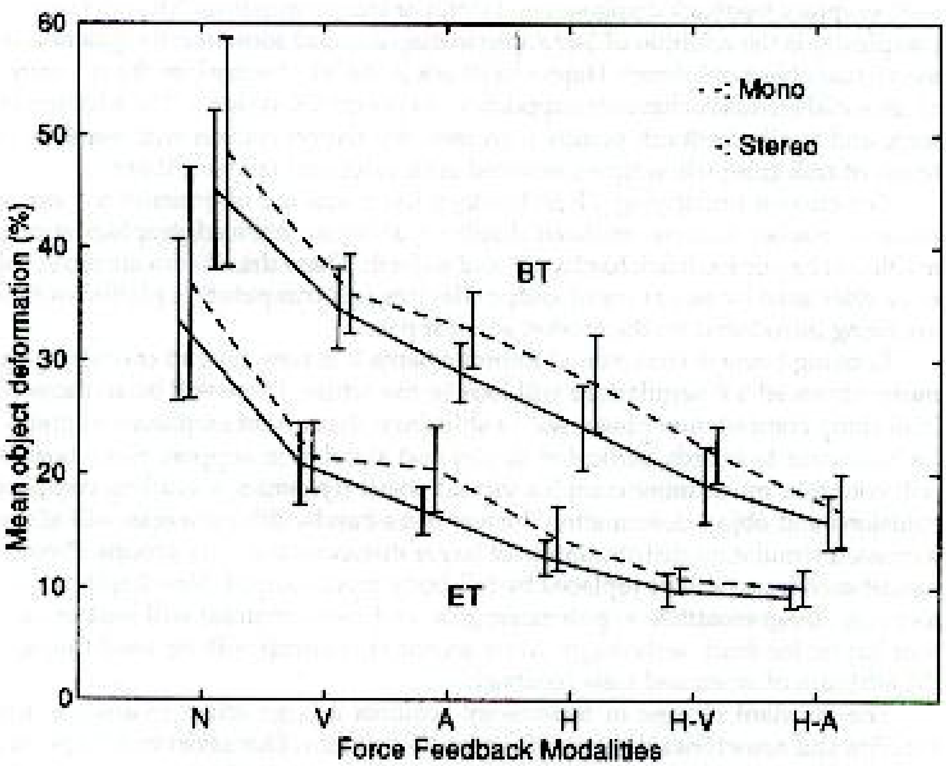

Figure 12. Dextrous manipulation task with various feedback modalities: beginning of trial (BT) and end of trial (ET; Richard et al., 1996). Reprinted by permission. Copyright 1996, Massachusetts Institute of Technology.

Very recent trials have confirmed the importance of reducing friction and inertia in haptic devices used in VR simulations. These tests replaced the RM I used in the human-factor study described earlier with an improved "Rutgers Master II" (RM II). This force feedback device has custom, low-friction glass/graphite cylinders and colocated sensors and actuators. It was thus possible to reduce static friction by an order of magnitude compared with the RM I. When the dextrous manipulation task protocol developed by Richard et al. (1996) was repeated with the RM II (Fabiani, Burdea, Langrana, & Gomez, in press), the mean virtual ball deformation and standard deviation were smaller. Thus, the reduced friction and increased dynamic range resulted in a more responsive input–output device.

CONCLUSIONS

"Traditional" VR systems have integrated position trackers and sensing gloves with graphics feedback displayed on HMDs or simple monitors. The next step in complexity is the addition of 3-D audio to map a sound source on the position of the virtual object of interest. Haptic feedback is the latest entry into the spectrum of sensorial feedback channels supported by current VR systems. The addition of force and tactile feedback greatly increased simulation realism with benefits in terms of task completion times, reduced error rates and learning times.

The current underlying VR technology has a number of limitations such as reduced tracker volume, reduced display resolution, reduced graphics speed, inefficient haptic feedback hardware, and so forth. These drawbacks are expected to be alleviated by newer input–output devices and computation platforms that are being introduced on the market at a fast pace.

Looking beyond incremental improvements it is now time to consider how more advanced VR simulations will look in the future. There will be an increase in desktop computation power with a shift from the current emphasis on graphics hardware to boards dedicated to physical simulation support. Such boards will compute much more complex virtual object dynamics, including complex collisions and object deformation. Newer large-bandwidth networks will allow increased simulation distribution over larger distances and user groups. Present hand-based input will be replaced by full-body input–output. New haptic actuators (e.g., magnetostrictive, polymeric gels, and piezo motors) will replace current haptic feedback technology. More sensorial channels will be used through the addition of smell and taste feedback.

The constant change in technology requires a large effort to analyze the benefits and sometimes dangers of newer VR systems. Our effort in VR-specific human-factors research has not kept pace with the technology evolution. It is thus necessary to intensify efforts in the study of human–virtual environment interaction and then use these results to improve technology, and fine tune its applications.

REFERENCES

Akka, R. (1992). *Utilizing 6D head-tracking data for stereoscopic computer graphics perspective transformations*. San Rafael, CA: StereoGraphics Corp.

Bardot, I., Bochereau, L., Bourgine, P., Heyd, B., Hossenlopp, J., Martin, N., Rogeaux, M., & Trystram, G. (1992). Cuisiner artificial: Un automate pour la formulation sensorielle de produits alimentaires [Artificial oven: A robot for synthesizing the smells of food]. *Proceedings of Interface to Real and Virtual Worlds Conference*, 451–461.

Bergamasco, M., Allotta, B., Bosio, L., Ferretti, L., Parrini, G., Prisco, G., Salsedo, F., & Sartini, M. (1994). An arm exoskeleton system for teleoperation and virtual environments applications. *Proceedings of the IEEE International Conference on Robotics and Automation*, 2, 1449–1454.

Bowman, S. (1993). Global devices 3D controller. *Real Time Graphics*, 1(8), 8.

Burdea, G. (in press). *Force and touch feedback for virtual reality*. New York: Wiley.

Burdea, G., & Coiffet, P. (1994). *Virtual reality technology*. New York: Wiley.

Burdea, G., & Gomez, D. (1994). *Actuator system for providing feedback to a portable master support*. U.S. Patent No. 5,354,162.

Burdea, G., Zhuang, J., Roskos, E., Silver, D., & Langrana, N. (1992). A portable dextrous master with force feedback. *Presence—Teleoperators and Virtual Environments*, 1, 18–27.

Codella, C., Jalili, R., Koved, L., & Lewis, B. (1993). *A toolkit for developing multi-user, distributed virtual environments*. Yorktown Heights, NY: IBM T. J. Watson Research Center.

Cutkosky, M., & Howe, R. (1990). Human grasp choice and robotic grasp analysis. In S. Venkataraman & T. Iberall (Eds.), *Dextrous robot hands*. New York: Springer-Verlag.

Cutt, P. (1993). Tactile displays: Adding the sense of touch to virtual environments. *Proceedings of Virtual Reality Systems '93 Conference*, 120–122.

Elmer-Dewitt, P. (1993, February 8). Cyberpunk. *Time Magazine*, 58–65.

Exos Inc. (1993). *The touch master* [Company brochure]. Woburn, MA: Author.

Fabiani, L., Burdea, G., Langrana, N., & Gomez, D. (in press). Human performance using the Rutgers Master II force feedback interface. *IEEE Virtual Reality Annual International Symposium*.

Faris, S. (1992). True stereo display and hard copy from one system. *Advanced Imaging*, 18–22.

Foley, J., van Dam, A., Feiner, S., & Hughes, J. (1990). *Computer graphics: Principles and practice*. Reading, MA: Addison-Wesley.

Gomez, D., Burdea, G., & Langrana, N. (1994). The second-generation Rutgers Master (RM-II). *The Third International Conference on Automation Technology*, 5, 7–10.

Gouraud, H. (1971). Continuous shading of curved surfaces. *IEEE Transactions on Computers, c-20*, 623–629.

Hirzinger, G. (1987). The space and telerobotic concepts of DFVLR Rotex. *Proceedings 1987 IEEE International Conference on Robotics and Automation*, 443–449.

Immersion Co. (1993). *Immersion probe*. Palo Alto, CA: Author.

Keller, P., Kouzes, R., Kangas, L., & Hashem, S. (1995). Transmission of olfactory information in telemedicine. In K. Morgan, R. Satava, H. Sieburg, R. Mattheus, & J. Christensen (Eds.), *Interactive technology and the new paradigm for healthcare*. Amsterdam: IOS.

Kennedy, R. S., & Stanney, K. M. (1996/this issue). Postural instability induced by virtual reality exposure: Development of a certification protocol. *International Journal of Human–Computer Interaction*, 8, 25–47.

Kramer, J., Lindener, P., & George, W. (1991). *Communication system for deaf, deaf–blind, or non-vocal individuals using instrumented glove*. U.S. Patent No. 5,047,952.

Krieg, J. (1993). Motion tracking: Polhemus technology. *Virtual Reality Systems, 1*, 32–36.
Krueger, M. (1991). *Artificial reality II*. New York: Addison-Wesley.
Marcus, B. (1993, October). *Sensing, perception and feedback for VR*. VR Systems Fall 93 Conference, New York City.
Marcus, B., Lawrence, T., & Churchill, P. (1991). *Hand position/measurement control system*. U.S. Patent No. 4,986,280.
Massie, T., & Salisbury, K. (1994). The PHANToM Haptic Interface: A device for probing virtual objects. *ASME Winter Annual Meeting, DSC-55*, 295–300.
Moshell, J. M., & Hughes, C. E. (1996/this issue). The virtual academy: A simulated environment for constructionist learning. *International Journal of Human–Computer Interaction, 8*, 95–110.
National Research Council. (1995). Haptic Interfaces. In N. Durlach & A. Mavor (Eds.), *Virtual reality: Scientific and technological challenges*. Washington, DC: National Academy Press.
Pimentel, K., & Teixeira, K. (1993). *Virtual reality: Through the new looking glass*. New York: Windcrest/McGraw-Hill.
Richard, P., Burdea, G., Birebent, G., Gomez, D., Langrana N., & Coiffet, P. (1996). Effect of frame rate and force feedback on virtual object manipulation. *Presence—Teleoperators and Virtual Environments, 5*, 1–14.
Robinett, W., & Rolland, J. (1992). A computational model for the stereoscopic optics of a head-mounted display. *Presence—Teleoperators and Virtual Environments, 1*, 45–62.
Roehl, B. (1995a, May–June). The fifth dimension glove. *VR World*, 64–65.
Roehl, B. (1995b, May–June). The Virtual i/o "i-glasses!" HMD. *VR World*, 66–67.
RPI Advanced Technical Group. (1993). *Product catalog*. San Francisco: Author.
Scully, J. (1993). Motion tracking: The Ascension Flock of Birds. *Virtual Reality Systems, 1*, 40–42.
Sense8 Co. (1995). *WorldToolKit reference manual Version 2.1b*. Sausalito, CA: Author.
Sheridan, T. (1992). Defining our terms. *Presence—Teleoperators and Virtual Environments, 1*, 272–274.
Sowizral, H., & Barnes, J. (1993). Tracking position and orientation in a large volume. *Proceedings of IEEE Virtual Reality Annual International Symposium*, 132–139.
Spea Video Seven. (1992). *Fire RISC Graphics Board* [Company brochure]. Thame, UK: Author.
Wenzel, E. (1992). Localization in virtual acoustic display. *Presence—Teleoperators and Virtual Environments, 1*, 80–107.