CHAPTER 1

Introduction to Virtual Reality

1.1 WHAT IS VIRTUAL REALITY?

When we speak of "virtual reality" (VR) we refer to a computer simulation that creates an image of a world that appears to our senses in much the same way we perceive the real world, or "physical" reality. In order to convince the brain that the synthetic world is authentic, the computer simulation monitors the movements of the participant and adjusts the sensory display or displays in a manner that gives the feeling of being immersed or being present in the simulation. Concisely, virtual reality is a means of letting participants physically engage in some simulated environment that is distinct from their physical reality.

Virtual reality is a medium, a means by which humans can share ideas and experiences. We use the word *experience* to convey an entire virtual reality participation session. The part of the experience that is "the world" witnessed by the participant and with which they interact is referred to as the *virtual world*. However, the term "virtual world" does not only refer specifically to virtual reality worlds. It can also be used to refer to the content of other media, such as novels, movies, and other communication conventions.

Here is a more formal definition for virtual reality from Sherman and Craig:

> *A medium composed of interactive computer simulations that sense the participant's position and actions, providing synthetic feedback to one or more senses, giving the feeling of being immersed or being present in the simulation.*

Note that the definition states that a virtual reality experience provides synthetic stimuli to one or more of the user's senses. A typical VR system will substitute at least the visual stimuli, with aural stimuli also frequently provided. A third, less common sense that is included is skin-sensation and force feedback, which is jointly referred to as the *haptic* (touch) sense. Less frequently used senses include *vestibular* (balance), *olfaction* (smell), and *gustation* (taste).

FIGURE 1-1

A virtual reality participant wearing a head-mounted display and a glove input device interacts with a virtual world.

Image courtesy NCSA

There are many specialty hardware devices involved in bringing the rendered sensory images to the user from the proper perspective. A familiar VR visual display device is the head-mounted display (HMD). An HMD is a device that the user wears on the head, containing a screen positioned in front of each eye. Another common technology used to display the visual part of a VR experience is to project the images onto a large screen or multiple screens that cover a sizable amount of the participant's view. Such displays date back to flight simulation projection domes and to the work of Myron Krueger (an early VR researcher) in the 1970s. This type of VR visual display is generically referred to as a large-screen stationary display.

As our formal definition suggests, an equally if not more important aspect of a virtual reality system is sensing the participant's position. Without information about the direction the user is looking, reaching, pointing, etc., it is impossible for the VR output displays to appropriately stimulate the senses. Monitoring the user's body movements is called *tracking*.

There are some related technological terms that are also often used in the discourse of virtual reality technology. However, these terms are not necessarily restricted to VR. One such term is "cyberspace." *Cyberspace* is the notion that people who are physically located in disparate physical locations can, through the use of some mediating technology, interact as if they were physically proximate. Thus, even technology such as the telephone can put two or more people in the same cyberspace.

Two other terms related to virtual reality and to one another are "telepresence" and "augmented reality" (AR). *Telepresence* is similar to VR, in that it is a means to virtually place a participant in another location in which they are not physically present. The difference from VR is that this location is actually a real place that for one reason or another is too difficult, dangerous or inconvenient for the person to visit in person. Like telepresence, augmented

FIGURE 1-2

Though often misperceptions surround the technology, virtual reality holds promise for a wide range of present and future applications.

FIGURE 1-3

Applications of virtual reality range from medicine to science to entertainment. Recent advances allow developers to port commercial computer programs to VR systems with relative ease.

Image courtesy of Jeffrey Jacobson

reality gives the user an altered view of the real world. However, the view they are given is of their current physical location, and using technology with many characteristics in common with virtual reality, additional (virtual) information is added to their normal sensory input. Frequently, it is the visual sense that is augmented, providing the user with abilities such as peering through walls, or into a patient's body.

1.2 THE BEGINNINGS OF VR

If one considers virtual reality to be the simulation of an environment that allows a person to experience some place and event other than where they actually are and what is actually happening around them, then flight simulators are an early example of this medium. Flight simulators based on interactive computer displays date back to the early 1970s. Earlier flight simulators made use of mechanically driven instrument displays driven by linkages to the pilot's flight controls such as the yoke, rudder pedals, etc. Many of the precomputer flight simulators were pedantic mechanical devices to give a future pilot the opportunity to become familiar with the flight controls and displays.

Later, by controlling the motion of a video camera over a scale model of some terrain, a sense of immersion was created. Although this did fulfill the criteria for virtual reality portrayed in the opening paragraph of this section, these early flight simulators were not general-purpose environments. A different simulator must be constructed for each type of aircraft, and additional terrain models created for new locations. General-purpose simulation was only possible after the advent of advanced computer graphics and display technologies.

In the following 11 examples of research efforts of different groups in VR development one can gain a sense of how VR technology came to be.

1.2.1 Morton Heilig's Sensorama

Early sensory display experiences included the Sensorama. The Sensorama was the brainchild of cinematographer and inventor Morton Heilig. Demonstrated in 1956, Sensorama was a scripted multimodal experience in which a participant was seated in front of a display screen equipped with a variety of sensory stimulators. These stimulator displays included sound, wind, smell, and vibration. The noninteractive scenario was driving a motorcycle through an environment with the appropriate stimulators triggered at the appropriate time. For example, riding near a bus exposed the rider to a whiff of exhaust.

The Sensorama system, however, was lacking a major component of the modern virtual reality system: response based on user's actions.

1.2.2 Ivan Sutherland's vision for computer-based virtual reality

In 1963, Harvard graduate student Ivan Sutherland demonstrated Sketchpad, a system to allow interactive, computer-generated visual imagery displayed on a cathode ray tube. In 1965, he described a vision for an immersive, computer-based, synthetic-world display system. His vision included the presentation of visual, aural, and haptic feedback in appropriate response to the user's actions. By 1968, Sutherland (as a professor at the University of Utah) had realized and publicly demonstrated a system that accomplished the visual component of his vision.

Sutherland's system included an HMD, mechanical head tracking using spooled retractable cables, and a computer program that rendered a simple stick representation of a cyclo-hexane molecule in three dimensions.

Sutherland later cofounded Evans and Sutherland Computer Corporation (E&S) and developed sophisticated real-time graphics rendering hardware for the flight simulator community.

1.2.3 Myron Krueger's Videoplace

Following Sutherland's demonstration, a variety of research and development efforts were born in university laboratories, government and military facilities, and, later, in the commercial sector.

In the academic community, University of Wisconsin researcher Myron Krueger was experimenting with a different perspective on virtual reality systems, which he referred to as "*Artificial Reality*." Whereas Sutherland's head-mounted display was especially suited for a first-person point of view in the virtual world, Krueger's artificial reality provided a second-person view of a virtual world in which participants could watch themselves within the world.

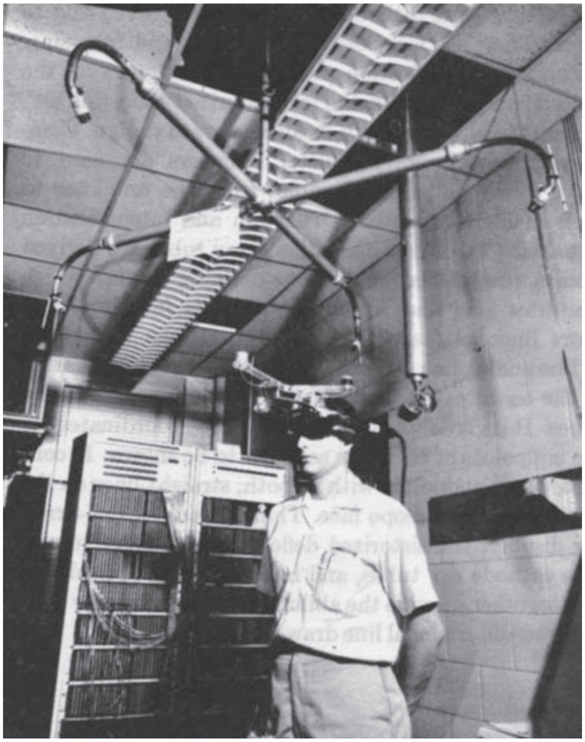

FIGURE 1-4

Ivan Sutherland demonstrates the first HMD virtual reality system.

Image courtesy Ivan Sutherland

Krueger's systems also differed from Sutherland's work in that he used video camera inputs to track the user's movements. Use of video camera technology resulted in two significant differences: The machine's perspective of the user was from the second-person point of view, and the user was not encumbered by any mechanical devices or other sensors attached to their body.

Other universities pursued various aspects of the virtual reality problem.

1.2.4 University of North Carolina at Chapel Hill

In the late 1960s, the University of North Carolina at Chapel Hill (UNC) computer science department founder and professor Fred Brooks espoused the need to have development work geared toward specific application problems. For example, a chemist would be interested in how two molecules dock together. Brooks' team also measured the benefits and pitfalls of their various innovations.

Due to the unavailability of capable hardware at the time, UNC also had to focus on hardware development, including high-performance graphics engines, head-mounted displays, and a variety of input and output devices, including devices to provide haptic feedback in

FIGURE 1-5

In Krueger's Artificial Reality, *a video camera is used to place an overlay of the participant's body on the virtual world.*

Image courtesy Myron Krueger

the form of responsive forces. Several commercial products have evolved from the innovative research at UNC.

1.2.5 Electronic Visualization Lab at the University of Illinois at Chicago

At the University of Illinois at Chicago, Tom DeFanti and Dan Sandin cofounded the Electronic Visualization Lab (EVL), where different types of graphical representations, input and output devices, and interaction techniques were explored. Most notable among their achievements were the development of the Sayer glove in 1977 (a glove outfitted to sense the bend of the wearer's fingers) and, in 1992, the announcement of the CAVE™ visual display system. The CAVE is a walk-in virtual reality theater typically configured as a 10-foot cube with three or more of its surfaces rear-projected with stereoscopic, head-tracked, computer graphics.

1.3 VR PARADIGMS

While we have already mentioned that VR systems provide synthetic stimuli to the senses, it is important to note that there are multiple ways by which this can be accomplished. Many

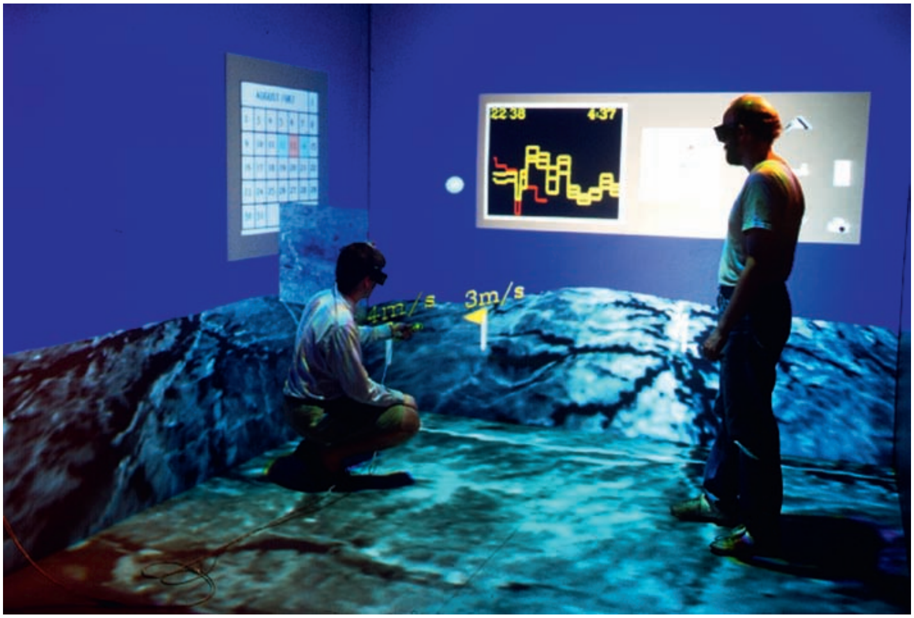

FIGURE 1-6

In the CAVE, participants stand surrounded by screens onto which the virtual world is displayed.
Image courtesy NCSA

suitable display technologies exist, but in general they can be categorized into three display paradigms. These three basic paradigms hold for not only visual displays, but also for display to other senses such as aural and touch (haptic) display systems. *Stationary* displays are fixed in place. Although the display doesn't move, the world is rendered in response to the user's bodily position. Examples of stationary visual displays include CAVE-type systems, single large screen systems, and desktop monitors. Loudspeakers are an example of stationary aural displays.

Head-based displays move in conjunction with the user's head. Consequently, no matter which way users turn their head, the displays move, remaining in a fixed position relative to the body's sensory inputs. Thus, visual screens remain in front of the users' eyes, and headphones on their ears. Examples of head-based visual displays include the helmet-type display often seen in popular media, and BOOM™-type displays which are a display box into which a user peers that can be moved around on mechanical linkages. Headphones are an example of head-based aural displays.

Hand-based displays are a special case of the head-based paradigm. In this case, users hold the display in their hand. For visual hand-based displays, monitoring both the user's head position as well as the position of the display is required, because the direction of view is important. Most often visual hand-based displays are used to overlay computer graphics imagery registered with the real world. An example of a haptic hand-based display is the

FIGURE 1-7

The stationary CAVE visual display and loudspeaker aural display are often used together.

Photography by William Sherman

FIGURE 1-8

The BOOM head-based display mounts the screens on an arm that keeps the weight from being applied to the user's head.

Image courtesy Fakespace Systems, Inc

SensAble Technologies PHANToM™ arm. The PHANToM provides a dual role by mechanically tracking the user's hand as well as providing a force display to the hand.

1.4 COLLABORATION

One of the strengths of virtual reality is its capability to transcend the barriers of time and space. This transcendence leads to VR being a good vehicle for supporting collaboration. VR environments can foster collaboration in a number of different ways. Space can be shared, either physically or virtually. Dialog can be held synchronously, or in an asynchronous form.

Large-screen stationary systems such as the CAVE are the best type of VR system for collaborating in the same physical space. Many participants have a concurrent view of the virtual world, allowing them to point out items of interest to one another.

Most forms of VR systems provide a good way to collaborate in the same virtual space.

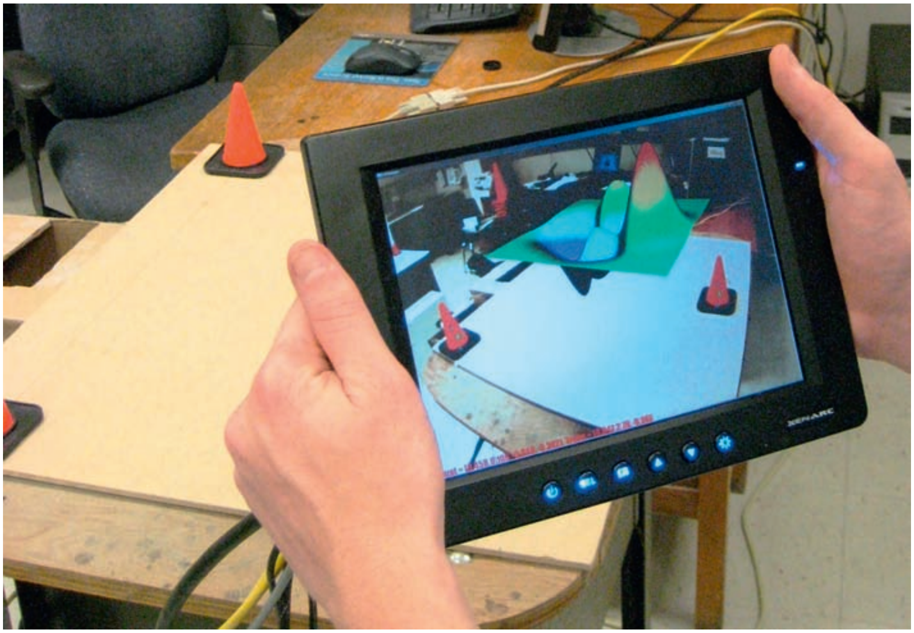

FIGURE 1-9

A hand-based display can provide an alternative view into the real or virtual world. In this example, researchers at the Colorado School of Mines use a hand-based display to superimpose virtual models on a stationary scene.

Image courtesy of Tyrone Vincent

A major benefit of virtual shared spaces is that they allow collaboration to take place via computer networks. Thus, not only can two workers share a space while remaining in their offices just down the hallway from one another, but they can also be an ocean away.

When working in a networked collaborative environment, each participant can be represented as a virtual entity. A virtual entity that represents a human in a collaborative environment is called an *avatar*. An avatar may be a somewhat realistic representation of the person, or an abstract representation. The mere presence of avatars can greatly improve the ability of the collaborators to communicate through nonverbal means. For example, pointing in a direction, waving an arm, or even just looking in a certain direction can convey valuable information from person to person.

Not every sense always needs to be transmitted in collaborative environments. For example, a telephone supports voice-only collaborations. VR, however, allows the option of participants sharing a three-dimensional world populated by 3-D objects that can be manipulated and worked with. Except for certain physical activities, most collaborative work relies only on the visual and aural senses, both of which are strengths of current VR technology.

Collaborators can inhabit the shared virtual world concurrently and engage in synchronous dialog and actions, or participate asynchronously by saving the state of the system after their component of the collaborative activity. Another possibility of asynchronous collaboration is to record all the actions of the participant(s), allowing other participants to replay that experience at a later time. In fact, the collaborators can leave annotations (such as messages or virtual pictures) for others who enter the space at a later time.

1.5 VIRTUAL REALITY SYSTEMS

The creation of a virtual reality system requires the integration of multiple components. These components include the system hardware, underlying support software for linking the display and input hardware together, the virtual world content with which the user will interact, and a user interface design that provides a suitable means for appropriate user interactions.

1.5.1 Hardware

Hardware used in virtual reality systems can be roughly categorized as display devices, input devices that a user consciously activates, and input devices that monitor the user, along with the computer that supports the modeling and rendering of the virtual world.

1.5.1.1 Computer/graphics engine

The main computing engine is responsible for calculating the physical behavior of the virtual world, and then rendering the state of the world into visual, aural, haptic, etc. representations. Because an effective VR experience requires real-time interactions, the computer system has some specific requirements.

The computational system can be implemented on a single large computer that meets all the requirements, or it can be implemented on multiple computers. In the latter case, the cadre

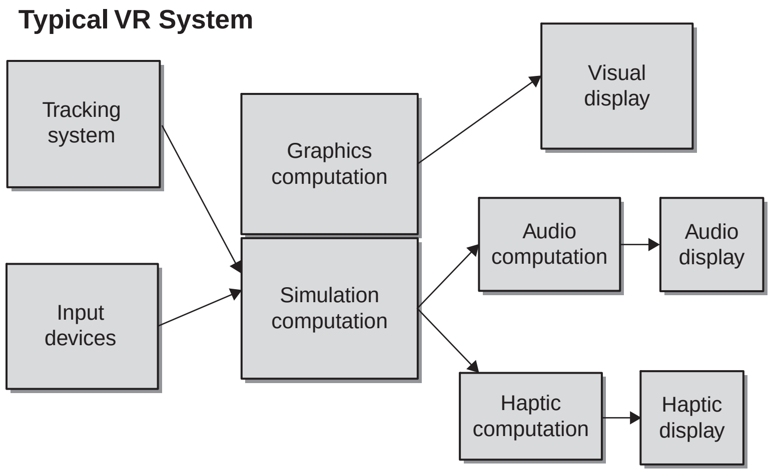

FIGURE 1-10

This diagram illustrates how the various components are integrated in a typical VR system.

of machines must be interconnected via a low-latency, high-speed communication network. Latency (the time delay between the time an event occurs and the time its results are apparent) is an important factor in any VR system. Any latency in the overall system reduces the effectiveness of the system. The use of multiple CPU components allows the system to achieve more computations both for the graphics and for the world simulation.

The primary needs that the computing system must meet include enough computational power to perform the virtual world's physics simulation calculations, sufficient graphical rendering performance from a "graphics engine" computational component, a means of rendering sounds, and perhaps rendering of other senses such as haptic (touch) information.

The specific computational needs vary based on the type of applications the system will be required to run. Representations of the real world generally require the ability to map pictures of the world onto surfaces to deliver a look of high detail (texture maps); however, an application to visualize a molecule could be done without the use of such features, requiring instead an increased geometric throughput. Some worlds, those that consist only of static objects, require no computation for world physics.

In addition to rapidly rendering the graphical representations, the graphics engine should have the capability to synchronize the display updates between multiple displays for rendering to both eyes (stereoscopic vision) in a head-based display, to multiple screens on a multi-screened projection display, or perhaps to multiple-projectors projecting overlapping left- and right-eye images to the same screen. Many high-end graphics engines and projectors have provisions to render and display stereoscopic images through a single mechanism. The absence of synchronization between displays leads to negative artifacts such as the world appearing discontinuous between two neighboring screens.

Modern computing systems include the ability to render sounds. If more advanced aural rendering operations are required, however, signals can be sent to an external audio processor. For example, the ability to make sound appear to come from a particular location relative to the user's head (spatialization) generally requires additional audio-rendering hardware.

The ability to perform multiple operations at the same time is also an overall requirement of VR systems. Thus, having an operating system capable of true multithreaded operation is a prerequisite of VR systems. The use of multiple computers is also a means of accomplishing this need.

Through the 1990s, many large VR projects relied on the use of larger (refrigerator-sized) computer workstations with multiple CPUs, and multiple instances of high-performance graphics-rendering hardware. However, with the advent of 3D graphics accelerators aimed at the consumer game market, it is now possible to utilize personal computers to implement VR systems with virtual worlds of significant complexity. This "low-cost VR" has made virtual reality systems available to a whole new class of users by decreasing the cost by an order of magnitude or more. A notable example of makers of such systems is Visbox incorporated, whose VisBox systems currently cost under $100,000.

1.5.1.2 Visual displays

The visual display portion of a virtual reality display generally has the most influence on the overall design on the virtual reality system. This influence is due to the visual system being the predominate means of communication for most people. It also tends to dominate how a VR system is defined, including which display paradigm is implemented.

Each type of visual display paradigm (stationary, head-based, and hand-based) has its own specific benefits and disadvantages, which are further influenced by advances in technology, and the amount of monetary resources available. In addition to these basic paradigms, all the visual displays can either display stereoscopic images, or monoscopic. In general, because virtual reality often attempts to mimic the sensation of physical reality, stereoscopic display is presumed.

Large-screen stationary displays such as the CAVE, wall displays, and table or desk displays use fixed-position screens to fill a relatively large portion of the field-of-view (FOV) for one or more viewers. Many of these displays (such as the CAVE and CAVE-like systems) wrap screens around the participant, surrounding the user as much as possible with the visual representation of the world. Even systems with a single display surface can fill significant portions of the user's view when the user stands near the large screen.

Thus, a primary benefit of the large-screen stationary display is FOV coverage. Other advantages include the reduced amount of hardware worn by users, which improves the ability to see colleagues physically standing next to them due to the reduced negative impact of latency. The ability of the user to continue to see the physical world while viewing the virtual world also improves the safety of the system.

Downsides of this style of visual display potentially include an incomplete view of the virtual world (field-of-regard), cost, and the difficulty of masking the real world if desired. The cost of these displays can vary greatly depending on the degree to which the user is surrounded and whether multiple projectors are used to increase the resolution of the imagery by tiling them together. An increased number of projectors also means more graphics-rendering hardware will be needed. Currently, with the use of projected images, the amount of space required is also one of the costs of using a large-screen display. The limited field-of-regard problem is solvable with an added cost. Six-sided (cube) CAVE-like facilities have been built that entirely surround the participant with screens (one being a door). The cost of such a facility needs to include creating a surface on which multiple people can stand, while being projected onto from below. This has been a significant challenge in the development of such systems.

Head-based displays (HBD) are perhaps the most commonly thought-of type of virtual reality display, having been popularized in movies and television. Early forms of head-based displays were often mounted onto fighter-pilot helmets, and thus were referred to as helmet-mounted displays (HMDs). Later the acronym HMD was also used for "head-mounted display." Either way, these devices were typically heavy headsets with attached screens positioned in front of the wearer's eyes. Two other types of HBDs that have also become available are a mechanical arm-mounted display that users pull up to their face, without any weight being placed on their body. The original of this class, the Binocular Omni-Orientation Monitor (BOOM), counterbalances the arm with the display. Later versions of HBDs use smaller screens, and weigh significantly less than the original HMDs. As these displays become closer to the sensation of wearing a basic pair of sunglasses, there is an increased tendency to label them "head-worn displays," with the superior connotations that phrase implies.

A major benefit of the head-based display is that users can turn their head to see any direction in the world. This is called 100% field-of-regard. Other benefits include being generally cheaper than large-screen displays, requiring less space, and being much more portable.

A significant disadvantage of HMDs is that any latency in the VR system is more noticeable to the user and thus more likely to cause nausea or a headache (thus limiting the interaction time). The more widely used head-mounted displays have the problem of the additional weight that users must carry on their head, along with cables to carry the video and tracking information. BOOM and BOOM-like display armatures often extend to the floor. Thus, the armature frequently causes blind-spots to which the user cannot move. Also, while BOOMs do not put the weight on the user's head, the display has a certain amount of inertia that can affect the experience. Head-worn displays therefore sound like an optimal solution, but they typically have screens with much lower resolution than what can be provided in BOOMs, HMDs, and stationary displays. Another disadvantage of head-based displays is that they are limited to a single user at a time, have a narrower field of view, and generally isolate that user from the people around the user, making it hard to discuss an ongoing experience.

Desktop VR displays (also known as *fishtank VR*) are similar to the large-screen displays in that they fall into the stationary display paradigm. The popular term "fishtank VR" is derived from the way one peers into a desktop VR display. A desktop VR display is basically a standard computer monitor, often augmented with the ability to display stereographically. By combining the monitor with the necessary tracking and other input devices and VR software, the scene appears to actually be inside the display—the way fish are inside an aquarium. Thus, if viewers moves their head left or right, they can see the fish from a different perspective, and similarly for the objects in the virtual world.

FIGURE 1-11

A computer monitor with a video camera can be a very simple VR display referred to as "fishtank VR" due to the similarity with looking into an aquarium.
Photography by William Sherman

The major advantage of the desktop VR display is that it can usually make use of an existing desktop computer with a few inexpensive additions. Thus, the cost of creating such a system is not excessive. Another significant benefit is that it can be used right at the user's desk. Frequently, the more difficult it is to use a VR system, the less often it will be used, and going to another room or building to make use of the system requires that the user expect significant improvements in the experience above what a monitor, keyboard, and mouse can provide. In fact, computer hardware has progressed to the point where, with the addition of a camera for user tracking, a VR system can nearly be completely implemented on a laptop computer, except not many laptops offer stereoscopic display.

There are some big disadvantages to the desktop display. These include very limited field-of-view and very limited field-of-regard. Users are only able to see what is immediately in front of them, and a little off to the side when they lean over, like looking through a window. Compared with the other types of visual VR displays, the cost is minimal, but there are costs to upgrade to a stereoscopic image, along with some input hardware and software to track the user's movement. The best tracking solution has been to use a video camera.

Hand-based VR displays are have not been widely used by VR systems. When used, they typically have a specific VR experience that makes use of them and generally have a specific need that must be fulfilled. The most intuitive type of hand-based display is a pair of binoculars that contain two small screens instead of the typical lenses. The binoculars continuously display a magnified (computer processed) view in the direction they are pointed, and when the user holds them up he or she can see the processed image. Another style of hand-based display is to hold a screen approximately the size of one's palm in the hand. The image on this screen shows the virtual world from the perspective of where the user's eyes are through the small window. This form of display works well as a "magic lens" display, giving the user an altered view of the "reality." The altered view might operate as if it were an "x-ray vision" device. The "reality" that is altered can be either physical reality or a virtual object itself. The palm-sized screen form of display is typically displayed monoscopically, in part because it is difficult to acquire small flat screens that can display stereo images. Modern cellular "smart phones" are now powerful enough to be used in VR, and more frequently in AR applications.

Although not widely used, handheld VR displays do have some advantages. In particular, they have an advantage when a VR experience has a natural interface for which the display is perfectly suited—as with the binocular, or "magic-lens" interfaces described. Because the user can choose when to look at a handheld display, it can be combined with either physical reality (as an augmented reality display), or in a screen-based virtual reality display such as a CAVE. Thus, a virtual reality world can be augmented—*augmented virtual reality.* Another nice feature of hand-based displays is that they tend to not be very encumbering.

Where hand-based displays do not work well is when there is no other VR display and the application requires a reasonable amount of FOV. Both the binocular and palm-sized devices provide very limited FOV. And while the field-of-regard is technically 100%, it requires the user to move the device through a large spherical motion to see in all directions. In general, handheld displays are less immersive, except when used to augment a larger view (real or virtual).

1.5.1.3 Aural displays

The inclusion of an aural display in a virtual reality system is generally a good way of enhancing any experience for a minimal additional cost. Unlike the visual display, it cannot be assumed that the aural image is presented stereophonically. In fact, the notion of "stereo" is more complicated with the aural sense.

Many virtual reality experiences can utilize a single (monophonic) channel of sound and still provide a deeply immersive experience. Experiences that provide just an ambient background sound, perhaps combined with some discrete sounds that mark an event in the world, seldom require more than monophonic. When this isn't the case, the question becomes whether traditional stereophonic sounds should be used versus a more complex method of sound spatialization.

The trouble with traditional (prerendered) stereophonic sound display is not that it only comes out of two speakers, but rather that it is preproduced (prerendered) to seem as though particular sounds come from particular locations. Because virtual reality is interactive, it is not generally possible to know a priori where the sound will be relative to the listener. Thus, sounds that must appear to emanate from a particular location need to be processed to create this effect. The processing is referred to as *spatialization*. Spatialized sound can be rendered to function in two-speaker (binaural) or multispeaker displays.

An interesting discovery regarding spatialized sound is that it can be effectively combined with prerendered stereo and monophonic sounds. For example, a VR experience might have a sound associated with a particular object or person in the world. That sound therefore should be spatialized to seem as if it follows the object or person. The scene might also have generic street sounds in the background presented as prerendered stereo. A monophonic, ambient orchestration to influence the mood can be added to the mix to create an overall highly immersive effect.

The two common sound display devices are loudspeakers and headphones. These two styles match well with the stationary and head-based visual display paradigms respectively. *Loudspeakers*, the aural display of the stationary paradigm, work well with CAVE-like displays, large wall displays, and desktop displays. *Headphones*, the aural head-based display paradigm, work well with head-mounted, BOOM, and other head-based displays. Of course, it is also possible and sometimes desirable to use headphones in a CAVE, particularly if the sound spatialization system works best with them. Likewise, there are good reasons to use stationary speakers with a head-based system.

Often, a single subwoofer is added to output loud, low-frequency sounds. Only one subwoofer is required because low frequency sounds are not easily localized by the human auditory system.

The cost of most aural displays generally pales in comparison with the cost of the rest of the VR system. Thus, neither form of aural display is more advantageous in that respect. The primary advantages of the two systems are that loudspeakers can be more easily heard by a group of participants, and headphones are generally easier to use when producing spatialized sounds. Also, headphones have a slight safety disadvantage in that if an excessive signal is presented, it will be very close to the listener's ears.

1.5.1.4 Haptic displays

Roughly speaking, haptic displays relate to the sense of touch. However, not all of the haptic sensations come via the skin. Some of what is called "haptic display" is related to the muscular and skeletal systems. Therefore, haptic displays are generally discussed in the two component terms: "tactile" (input through the skin) and "proprioceptic" (input through the muscular and skeletal systems). Sensing the coarseness of sandpaper or the temperature of water are tactile sensations. Sensing how much effort is required to lift a box, or knowing the current location of one's arm are proprioceptic sensations.

Different technologies are generally required for creating forces versus creating subtle skin-response sensations. Therefore, most devices designed for haptic display focus on either tactile or proprioceptic presentation.

Like visual and aural display types, haptic displays can also be divided based on the stationary versus body-based paradigms. However, when discussing haptic displays, these characteristics are typically referred to as "world-grounded" (stationary) versus "self-grounded" (body-based) displays.

World-grounded displays are those that have a base attached to the ceiling or that perhaps sit on the desktop or are affixed in some way to some object in the real world. Typically, the user holds the end of an arm with multiple linkages leading back to the base. Each of the linkages is capable of exerting an active or resistive force in a particular direction. Thus, when the user grabs an object and tries to move it, the ease with which it can be moved can be felt, allowing the user to sense the weight of the object and the friction or viscosity of the containing medium. Or, if the object is animate, such as a dog, then grabbing it (or its collar) can lead to an active force felt by the user.

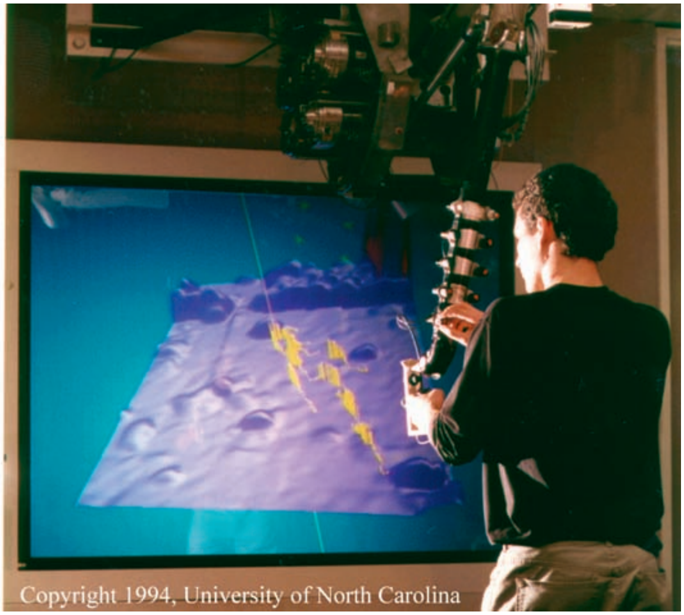

FIGURE 1-12

This world-grounded haptic force-feedback device is attached to the ceiling, allowing a user to grab the controls and interact with the molecular world.
Image courtesy the University of North Carolina at Chapel Hill

Self-grounded displays are those that are somehow worn by the user. A common example is a glove fitted with some form of tactile display, such as small vibrators. Force display devices can also be self-grounded, such as a display that resists

FIGURE 1-13

The Rutgers Dextrous Master II is a self-grounded display that can be used to prevent the fingers from closing all the way, simulating the effect of holding an object in the hand.

Image courtesy Rutgers University

the movement of the user's arm relative to their shoulder. The latter example works best however, either by ignoring the user's movement within the virtual world or by assuming that the shoulder is in a fixed location. In the latter case, the self-grounded arm display is effectively acting as a world-grounded display.

Another possible form of haptic feedback is that of the inactive prop device. In this case, the user gets tactile sensations from the skin touching a device and feeling its shape, texture, and sensing movement of buttons or other objects mounted on the prop. The prop device also provides some proprioceptic feedback by its weight and momentum. An example of an inactive prop is an instrumented (real) putter used as an interface to a virtual golf game.

1.5.1.5 Other sensory displays

Virtual reality systems make use of (in decreasing prevalence) visual, aural, and haptic displays. Use of other sensory displays has also been done. Of these the vestibular sense (the sense of balance) is the most common. In fact, it has been a very common form of display for flight simulation for decades. Olfactory display (smell) has been experimented with sparingly, and computer-controlled display of gustation (taste) is virtually nonexistent.

The most common form of vestibular display is the "motion platform." A motion platform is basically a large surface (the platform) mounted on top of hydraulic actuators that can raise, lower, and tilt the platform. The user (typically) sits on the platform, and in the case of flight simulators within a cockpit mounted to the platform. Sometimes the visual display is also mounted on the platform; other times it is projected onto a large dome that can be seen through the windows of the cockpit. By tilting the motion platform, the pilot can then sense when the aircraft begins to pitch, yaw, or roll, and by how much.

Another style of vestibular display is the bladder-equipped chair. By inflating and deflating different portions of the chair, the user can feel acceleration and deceleration. For example, when undergoing strong acceleration, pilots will feel themselves being pushed back in their chair. To recreate this, the bladder on the back of the display seat can be filled, and thus create a similar pressure sensation on the back of the pilot. A similar effect can be implemented for sensing the effective loss of gravity while riding in a roller coaster by deflating the seat of the chair.

1.5.1.6 Input devices and user tracking

Without input, a computer-generated display cannot be interactive, much less be considered a virtual reality system. In fact, virtual reality systems require not just a means for users to express their intentions, but also must track at least some subset of users' bodies. One can differentiate between these two types of input by referring to them as "cognitive input" (events specifically triggered by the user) and "user monitoring" (tracking the body movements of the user). Another way to think about this input dichotomy as an input that the user must specifically activate and an input that passively senses attributes, such as the position of the user.

The *position sensor* is the most important tracking device of any VR system. There are several types of position sensors, each with its own benefits and limitations. These sensors include electromagnetic, mechanical, optical, ultrasonic, inertial/gyroscopic, and neural/muscular devices. The most crucial factor of a position sensor is the type of limitations imposed on the system. Limitations generally arise from the technological means used to determine the relationship from some fixed origin and the sensor. For example, some trackers require an uninterrupted "line-of-sight" between a transmitter and a sensor. When the "line-of-sight" is interrupted (i.e., something comes between the transmitter and the sensor) the tracking system cannot function properly.

In position-sensing systems, there are three factors that play against one another (discounting cost): accuracy and precision of the reported sensor position, interfering media (e.g., metals, opaque objects), and encumbrance (wires, mechanical linkages). No available technology, at any cost, provides the optimal conditions in all three of these factors. Thus, the system designer must consider how the VR system will be used and make the optimal tradeoffs. One of the driving factors is simply the ability of the system to produce an acceptable experience. Noise and low accuracy in the position sensor reports, as well as high latency decrease the "realism" or immersiveness of the experience, and often can lead to nausea in some participants.

Electromagnetic tracking systems are popular input devices for VR systems because they do not require line of sight to the tracked object. However, because they use an electrically generated and received magnetic field to determine the six degrees-of-freedom of the sensor device, metals interfere with the functionality of such a system. Ferrous metals are particularly problematic. Also, active electronic devices in close proximity to a sensor can be an issue. The magnetic properties of metals within the VR environment cause distortions in where the sensor is perceived to be with respect to the transmitter. If the interfering metals are stationary, then minor distortions can sometimes be corrected for in software.

Fortunately, the amount of metal within the environment can often be controlled. Cases where particular care must be taken to improve tracking accuracy are head-worn gear made of metal or with internal electronics, and wheelchairs. In the case of HMDs or stereo glasses with electronics, the best solution is to locate the sensor as far away from the electronics as possible. In the case of a wheelchair, a sensor mounted to the participant's head is less of a concern than a handheld device that will be located closer to the metallic components of the chair.

Standard electromagnetic tracking systems have wires that connect with both the transmitter and the sensor units. This is somewhat encumbering, with cables tethering the participant to the VR system. For a greater cost, some of these systems connect the sensors, not directly to the VR system, but rather to a radio pack worn by the participant. The participant thus has more freedom to physically move about the space without the concern of tripping over wires.

Mechanical tracking systems use transcoders mounted on physical linkages to report the movement of the linkages. The position of the end point can be calculated from the transcoder values. The use of transcoders provides extremely accurate and precise position readings. By improving position reports, the overall VR experience is improved by giving an increased physically immersive sensation, and perhaps also reducing the likelihood of nausea. The overriding

FIGURE 1-14

A low-level electromagnetic field is emitted by the large black box. The signal is sensed by a receiving antenna, which allows the system to determine the location and orientation of the receiver.

Photography by William Sherman

problem of mechanical tracking systems is that there is some physical attachment between the user and the real world. This attachment can often impede the user from moving in a natural way. However, there are some situations where the user's movement is already restricted, and therefore the mechanical system can be designed such that no additional restrictions are added, such as a pilot sitting in a cockpit.

Glove input devices generally fall within the realm of mechanical position sensor. However, it is the configuration of the hand that is measured rather than the overall location and orientation of the entire hand. To deduce the shape of the hand, sensors are placed throughout the glove to determine the amount of bend between various joints. Two common bend-sensing technologies used for hand-position sensing are optical fibers that transmit less light when bent and metals that alter their resistance when bent.

Ultrasonic tracking systems use a collection of transducers—transmitters (speakers) and receivers (microphones)—that pass signals from one point to another. By measuring the time taken for the signal to arrive, one can compute (using the speed of sound) the distance between the transducer pair. The key factors in accomplishing a proper measurement are that multiple transducer-pair measurements are required to determine the complete (X, Y, Z, roll, pitch, and yaw) position, and an uninterrupted line-of-sight must be maintained between transducer pairs. Thus, hardware systems that use ultrasound to measure sensor positions typically mount several transducers on the sensor device to provide some redundancy, allowing the sensor to go through different orientations and still maintain sufficient contact with the transmitters.

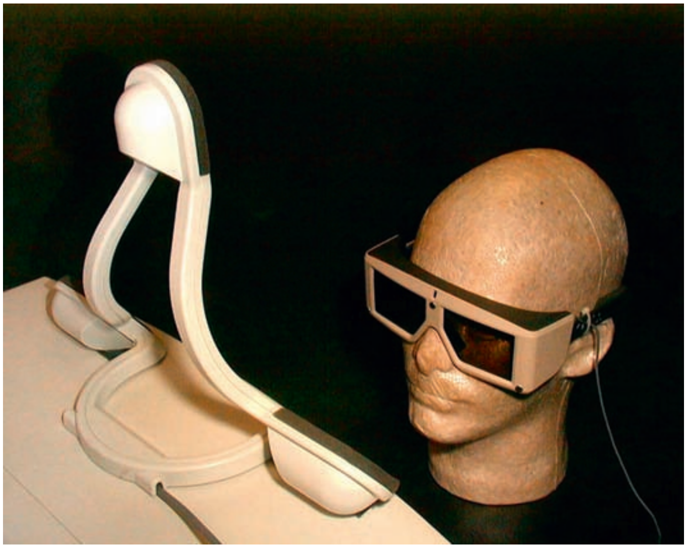

FIGURE 1-15

This basic ultrasonic tracking system uses three speakers and three microphones triangularly arranged to measure the distance between all the speaker-microphone pairs from which the location/orientation of the glasses can be determined.

Photography by William Sherman

Determining the orientation or location of a sensor requires that at least three transmitter-receiver transducer connections be made. In addition, there is a minimal distance that must exist between transducers in order to avoid ambiguous results. The number and spacing of the transducers can be cumbersome in some circumstances, such as adding significant weight to stereo-glasses, and requiring handheld devices to be large enough to accommodate the transducer distances.

Optical tracking systems can work along the same lines as ultrasonic systems, measuring distances by time and triangulation, or they may operate using computer vision by attempting to discern features of a video image to recognize where certain reference markers are located, and also how they are oriented. The markers used are generally designed to contrast with the rest of the scene. This contrast can be done by using illuminated objects such as light-emitting diodes, or by creating high-contrast signature shapes such as a white square surrounded by a black square.

Clearly, because the optical transducers work in the visual and near-visual spectrum, opaque objects will interfere with the operation of the sensors as there is a line-of-sight restriction with this form of tracking. However, optical tracking systems have some significant advantages over many other tracking systems. Specifically, a reasonable system can be constructed using commodity video equipment and freely available software. Another advantage is that video tracking can be done without the need for any wires emanating from the tracked sensor. One problem with video tracking is the reduced accuracy attainable using standard video resolutions.

Inertial and *gyroscopic* tracking systems are unlike many of the previously discussed methods in that they do not directly relate themselves to a fixed reference point. The downside is that they only report relative movements, not absolute positions. The benefit of this fact is that less hardware is required to implement these types of tracking. Thus, an inertial or gyroscopic tracker could be mounted in a small head-based display, and no other hardware would be required to give visually immersive feedback to the user. Another important benefit of this hardware is that there is very little lag between movement of the sensor and the reported movement.

The problem with such tracking systems is that because of the lack of a fixed reference, the reported values accumulate error. After a few minutes, when the user looks forward, the system may behave as if the user was looking ten or more degrees to the left or right. Frequently, inertial and gyroscopic tracking is combined with other tracking hardware so the benefits of each can complement one another. Because some VR systems, especially head-based systems, can cause nausea when there is significant lag between user movement and the visual image, the fast response of the inertial/gyroscopic system provides a low latency response to quick movements. Electromagnetic, ultrasonic, or other type of referenced tracking is then used to continually adjust for drift.

Neural and *muscular* tracking refer to the use of transducers placed on the skin of the participant to monitor muscular and other activity within the body, and make use of this information to provide inputs to the virtual reality system. For example, a sensor on the arm might be able to determine when the user clenches a fist. An example is a device called the Biomuse™. When the Biomuse is attached to the user's forearm, a virtual violin can respond to the user making bowing motions.

The tracking systems above are generally used to monitor the user's general body movements. This type of activity is referred to as that which is passively transmitted by the participant. Other VR input devices are designed to give the user more active and cognitive inputs. For example, pressing a button to jump forward in the virtual world is an active form of input.

Props and *platforms* are the physical places where such active input sensors are placed. The term "props" comes from the theater and film industry use of the word. Short for "property," a prop is any physical object that is not part of the scenery and can be manipulated by the actors. Thus in virtual reality systems a prop is an object that the participant can handle and use to interface with the virtual world. A prop may be embodied by a virtual object and might have physical controllers mounted on it.

Props themselves can be used for both passive and active user input. Handheld props are generally tracked in space, and thus a good indication of where one of the user's hands is located. Props also frequently have buttons, joysticks, and other input devices mounted on them, allowing the user to actively cause an action in the virtual world.

A platform is similarly used as a means of user input to the virtual world. It differs from the prop in that it is more like the scenery. Thus, a cockpit, or captain's wheel of a tall ship are both part of the "scene" where the participant is located and also provide a means of controlling the virtual world.

FIGURE 1-16

In this virtual reality system, a platform with a ship's wheel mounted on it provides the space where the participant takes a virtual voyage.

Image courtesy Randy Sprout

For gathering input about the real world for use in a virtual (or augmented) reality, there are many different types of data transducers. For example, MRI and medical ultrasonic scanners can produce data to recreate the internal organs of a patient. Real world objects and locations can be captured with laser scanning devices such as light detection and ranging (LIDAR). Larger scale locations and weather data can be collected by interpreting data transmitted by satellites.

Once the system has collected the input data it can further refine the data and otherwise filter it. Two common types of filters are for calibration/registration and gestures.

In order to provide a participant with a better sense of physical immersion, it behooves the system to respond in a manner consistent with the user's movement. In other words, if users move their head 4 inches, the system should not respond with an 8-inch movement. Many systems, and especially electromagnetic trackers, produce a consistent error that can be put in a table and used to compensate for the erroneous sensor reports. Other systems might be able to combine their data with fiducial (reference) markers that can be used to correct for slight errors in the data. Either method results in more accurate reports of sensor positions, at perhaps a slight increase in latency. Augmented reality systems are especially susceptible to poor calibration and registration to the real world because any errors are glaringly obvious against the real-world backdrop.

Another common form of filtering is to interpret patterns in the input from the user. For example, if the user extends both arms out the sides and repeatedly moves them up and down, the system may generate the "flap" input. Or if the system monitors finger movements and senses that the hand has closed into a fist, it may indicate to the virtual world that the user is attempting to grasp an object in the world. Sufficient tolerance must be built into a gesture recognition system to allow for variations from individual to individual.

Given the plethora of input possibilities, designers should consider the goals of the system and find the combination of input devices that best serve that goal.

1.5.2 Software

A variety of software components must be integrated to enable cogent VR experiences. Such software ranges from low-level libraries for simulating events, rendering display imagery, interfacing with I/O devices, and creating and altering object descriptions, to completely encapsulated "turnkey" systems that allow one to begin running an immersive experience with no programming effort.

1.5.2.1 Laws of nature—simulation code

Many VR experiences have some programmed laws of nature that govern the behaviors and interactions carried out by the objects in the world. The exception to this is the case where the only interaction possible is changing the user's viewpoint relative to the objects in the world. In this case, the user cannot manipulate the objects but only look at and work around them.

One option for "world simulation" is to merely allow several explicit cases of behavior to be executed under specific conditions. For example, in an architectural walk-through, the system

FIGURE 1-17

A simple, nonrealistic set of rules govern this fantasy space, providing both cartoon-like renderings as well as cartoon-like laws of nature.

Photography by William Sherman

may prevent users from walking through walls, and constrain users' vertical movement to be as if they were walking on the floor surfaces.

More advanced simulations can have global behaviors such as gravity, plus individual rules that apply only to specific objects. For example, a bee could be given a rule that allows it to fly in search of a flower, gather pollen, and then return to the hive. On the other hand, a flower could be plucked with a grasping gesture and when released, fall to the ground.

Other application simulations strive to more closely mimic the real world by adhering to mathematical descriptions of real physics. So, for a bee to fly, it would have to flap its wings sufficiently rapidly to generate the needed lift, and orient itself properly to send it in the desired direction.

Given that in a virtual reality experience there is no requirement that the world follow the laws of the real world, it is possible to give objects fantastic behaviors. Such behaviors might be to give the user "x-ray" vision abilities to see through objects or to see the interior structure of an object. Another possibility is to give the user the ability to move heavy objects such as walls and furniture in an architectural (or game-world) design application, and walk through walls.

The concept of world-physics also applies to how multiple users sharing the same world can interact and communicate with one another. Simple implementations of behaviors for collaboration might include representations of the other users (their "avatars"), and perhaps also an audio channel that allows everyone to communicate verbally.

1.5.2.2 Rendering libraries

Rendering libraries convert the form of the world from the internal computer database to what the user experiences. The rendering library must include the appropriate rendering algorithms for whatever sense is to be portrayed. Visual images produced from graphics rendering libraries are perhaps the most common of this class of software; however, such libraries have also been developed (and used in VR) for other senses, such as hearing and touch.

These libraries generally include features to render the basic elements of a "scene" along with features to enrich the display. For example, in a typical graphics library, along with the ability to render basic forms by specifying the vertices and colors of polygons, the programmer is also given options to add lighting elements, and overlay photographs onto polygons to make them appear more realistic. Also, such libraries can support higher level graphical functions like hierarchical object descriptions ("scenegraphs") and collision detection.

1.5.2.3 VR libraries

A complete virtual reality system is not comprised merely of rendering sensory outputs, but rather rendering appropriate outputs depending on the user's current position and actions. The paramount task of this VR library is to acquire the necessary information about the participant. This is done by interfacing with tracking and other input hardware. Information from the various sensors is integrated and provides the necessary parameters to the rendering systems. For example, the graphics (and also 3D audio) rendering process requires knowledge of the user's head position to give the proper visual/soundscape.

Another critical requirement of the library is to operate in "real time." It must perform all the input, simulation, and rendering functions at a rate that makes the world appear to be "real" by immediately responding to the participant's actions. Using multiple processing units on VR systems can help to achieve such "real time" responsiveness. Therefore, VR libraries typically include the ability to perform multiple tasks at once.

1.5.2.4 Ancillary software

The creation of a virtual reality experience also requires the use of various software in addition to the software required during the presentation of the experience. Examples of such tools include modeling software to aid in the construction of the form of the objects that inhabit the world; sound editing software to construct sound clips that will be heard in the experience, and image processing software to create appropriate texture maps.

Independent user interface libraries might also be linked with a VR experience to allow the operator to control parameters of the experience, for example a mouse-controlled widget panel. File formats such as VRML (a format for describing three-dimensional computer graphics objects and spaces) and other standard object formats also play an important role in the creation of virtual reality worlds.

1.6 REPRESENTATION

There are several stages to presenting information to the user. We have stated that virtual reality is indeed a medium for communication. As such, the choice of symbols one chooses to convey is important. Depending on the goal of the VR experience, one may choose to mimic the real world to a high degree of verisimilitude, or one can choose to disregard the structures and limitations extant in the real world and create surreal or fantastic worlds with never-seen-before objects, behaviors, and beings. One can choose to present aspects of real-world entities that are normally unseen, such as stresses within a structural beam, or present the world as perceived by someone who has undergone a traumatic brain injury.

Regardless of the application, a mapping must be made between concepts in the virtual world, and the stimuli that will be presented to the user's various sensory organs. When choosing representations for objects and concepts, tradeoffs must sometimes be made based on the limitations of the underlying systems and the requirements of the application. The choice of representations is clearly limited to the kinds of transducers available in the system. For example, most virtual reality systems provide a visual and aural display. Beyond these two modes of presentation, in some special cases there is extra hardware available for presenting certain tactile, force, smell, and vestibular feedback.

Within the modes of presentation, tradeoffs exist regarding fidelity versus cost and performance issues. For example a tradeoff in designing the visual aspects of an automobile lies between visual complexity/realism versus the real-time/interactive nature of the display. Limits on the real-time frame rate reduce the possible level of interactivity.

However, users who require highly complex extreme realism may be willing to accept the reduction in frame rate.

Often, specialized rendering tricks are used to increase the realism. This includes techniques such as texture mapping, level of detail (LOD) culling, and polygon decimation. Sometimes these tricks lead to a tradeoff between realism in the geometric form versus realism in the surface look. The technique of texture mapping photographic images onto a simple geometry is a common method of making a world look more realistic. However, the closer one approaches a texture-mapped object (especially when presented stereoscopically), the more apparent it becomes that the form is not a true representation of the object. The object looks like a cardboard cutout or stage background of a play rather than the actual entity.

As has been stated, users' avatars are their representation within the virtual world. There is a wide range in how one can create this personal representation. Perhaps the simplest is to restrict the avatar to a nonvisual, vocal presentation. In the realm of visual avatars, there are a variety of avatar options. The type of interpersonal communication required by the application affects the avatar representation requirements. If the capability of expressing nonverbal body language is required, then having a 3D model with articulated arms offers the ability to point, wave, and perform other gestures. If seeing the faces of other users is important to read their reactions to events, then an avatar comprised of a video representation becomes the preferred option. In a fantasy scenario, users may not want to accurately reflect their real-world counterpart at all.

1.7 USER INTERACTION

Virtual reality offers the opportunity for new modes of interaction not previously available with traditional computing systems. While offering new possibilities, a downside is that there is no established set of conventional idioms. Often interaction styles are borrowed from two-dimensional user interfaces. For example, pull-down menus can be imported into a three-dimensional virtual world.

Using borrowed idioms helps the user by providing a familiar means of interfacing with the computer. However, it may not take advantage of the potential richness of the 3D virtual environment. Even when using borrowed paradigms, questions still remain regarding where

FIGURE 1-18

The menu is an interface technique that has been adapted to the virtual reality medium from the realm of desktop computer interfaces.

Photography by William Sherman

to place them, which direction they should face, and other decisions that were obvious in the 2D worlds for which they were designed.

1.7.1 Interaction Techniques

If one starts with a blank slate, not considering previous 2D interface styles, then one can conceive of new interaction styles that can be broken down into four major categories.

The obvious mode of interaction in virtual reality is to mimic the actions required in physical reality. Thus, to move an object, a user can position their hand at the object's location, grasp it by closing their fingers, and then by moving their hand, change the position of the object. In the virtual world, this can be emulated by tracking the position of the hands and fingers. This is referred to as a *direct* form of interaction.

While *direct interaction* best mimics our methods of manipulating the real world, there are other ways in which we are accustomed to interacting with computers. These three other forms of interaction are referred to as *physical*, *virtual*, and *agent* interactions.

Physical interactions are those that are input to the virtual world through input devices that the user actually touches. In a conventional computer system, the most common physical inputs are through the mouse and keyboard. In a virtual reality system, devices such as a handheld wand, steering wheel, or glove input devices are examples of physical inputs.

FIGURE 1-19

Steering a vehicle is one way in which the participant can use a physical device to interface with the virtual world.

Photography by William Sherman

Virtual input interactions are ones in which the "devices" with which one interacts are a part of the virtual world itself. Thus, a virtual button is one that is rendered directly in the world and might be activated when the users hand comes in "virtual" contact with the button. Many virtual interactions rely on physical or direct interactions to activate the virtual device. So in the given example, a direct interaction is used to press the virtual button. An example in which a physical input is used to activate a virtual device is when a slider is rendered in the world (or just on the screen in a traditional desktop interface) to control a parameter such as volume. In both the VR experience and desktop metaphor, a physical button on the wand or mouse is pressed to manipulate the slider.

The fourth type of interaction is to express control parameters via an *agent*. In other words, by communicating with a computer entity (the agent), one lets their desires be known, and expects the system to comply. For example, to travel through a solar system world, one might say the name of a planet and be taken into orbit around the specified celestial object. In the real world, we might tell our chauffeur the name of the location to which we wish to travel, and expect to be taken there without any further input.

Having listed the four forms with which one can cognitively input information to the virtual reality system, it is appropriate to examine three broad categories of the types of interactions commonly performed in a virtual reality experience. These interaction categories are making selections, performing manipulations, and traveling.

FIGURE 1-20

Interacting with a graphical (virtual) controller is an example of a virtual input interaction, such as moving this table using a virtual slider.

Photography by William Sherman

1.7.2 Making selections

The primary selections one can make in a virtual world are selecting an object on which to act, or to select a direction in which to go.

There are a variety of ways of indicating a direction of interest. Many of these ways make use of the position of some part of the user's body, such as pointing with a finger, gazing with the eyes, or facing with the torso. One can also indicate direction with devices such as a joystick or steering wheel, or by referring to a coordinate system or some landmark-based reference system.

There are many natural ways in which a VR system designer might choose to allow the user to select an item in a virtual world. In some of the previous examples, the user makes contact with an item to activate it—making contact is one way to select an item. By making use of a selected direction, one can point to the object of interest. Through the use of voice recognition software, the user might just name the object, either from memory or from a menu listing possible selections. Or the VR system might provide a menu system that allows the user to point to the desired object or make contact with the object's name.

1.7.3 Manipulating the virtual world

Having selected an item, the user will often want to perform some manipulation on that item. In many cases, the process of selecting an item may be incorporated directly into the manipulation process. For example, moving a box might be performed by touching or pointing at the box, pressing a button, and then moving the hand that is making virtual contact with the box.

The manipulated element of the experience can be either an object of the virtual world or an attribute of the overall virtual reality system. For example, moving a car is manipulating an object of the world, whereas choosing a filename to store the current status of the world is an attribute of the virtual reality system.

There are two ways of acting on elements of the experience: in a way that mimics the action of forces on them, or by changing attributes of objects in the world or the system in supernatural ways. So, a car in the world can be changed from blue to red by applying virtual paint to the car (mimicking reality) or by selecting the new color from a menu (supernatural modification).

1.7.4 Navigation

Navigation describes how we move from place to place. In the real world, we navigate from place to place as we walk, drive, ski, fly, skate, and sail through the world. In a VR experience, there are several additional choices for how one might navigate through the environment.

For clarity, the term *navigation* can be divided into two subcomponents: *travel* and *wayfinding*. Travel is the act of controlling one's movement through the world, such as by physically walking or controlling an airplane yoke. Wayfinding is using information about the world to guide the direction and speed of travel.

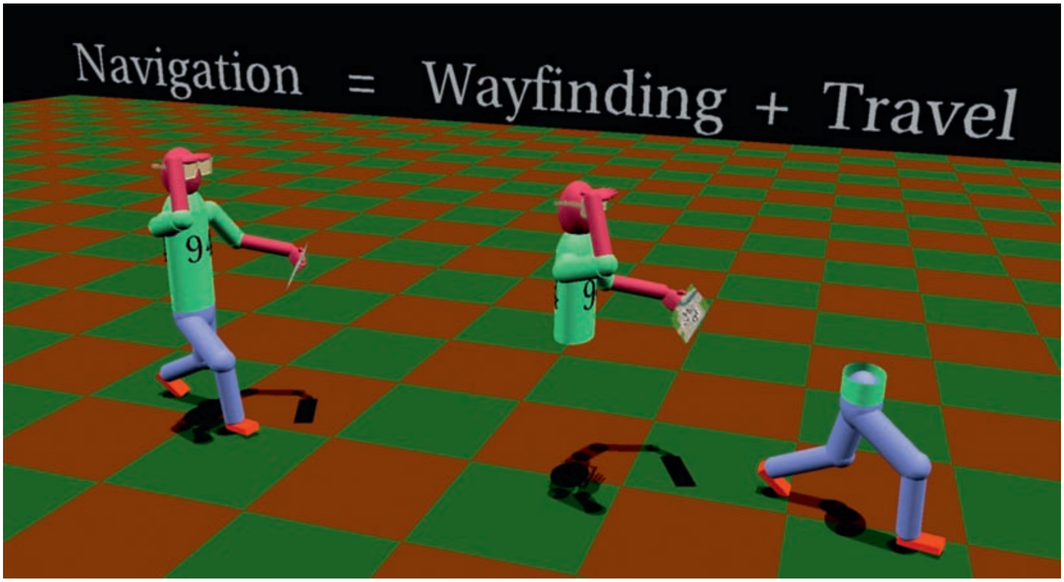

FIGURE 1-21

The task of navigating through a world can be broken into the component tasks of wayfinding (figuring out where you are and where to go) and travel (moving through the world).

Photography by William Sherman

There are 10 common travel paradigms used in virtual reality experiences:

- *Physical locomotion* is the simplest method of travel in VR. It is merely the ability for participants to move their bodies to change the position of their point of view within the virtual world. Physical locomotion travel is generally available in VR experiences, often in combination with another form of travel.
- *Ride-along* describes the method of travel that gives participants little or no freedom. They are taken along a predetermined path through the virtual world, perhaps with occasional choice-points. Usually participants can change their point of view or "look around" while on that path.
- *Tow-rope* travel is an extension of the ride-along paradigm. In this case, the user is being pulled along a predetermined path, but with the ability to move off the centerline of the path for a small distance.
- *Fly-through* travel is a generic term for methods that give the user almost complete freedom of control, in any direction. A subset of the fly-through method is the *walk-through*. In a walk-through interface, participants' movements are constrained to follow the terrain such that they are a natural "standing" height above it.
- *Pilot-through* describes the form of travel in which users controls their movements by using controls that mimic some form of vehicle in which they are riding.

- *Move-the-world* is a form of travel that is often less natural than the previous forms. Here, users "grab" the world and can bring it nearer, or move or orient it in any way by repositioning their hand.
- *Scale-the-world* travel is done by reducing the scale of the world, making a small movement, and then scaling the world back to its original size. The difference between the points about which the two scaling operations are performed causes the user to reappear at a new location when returning to the original scale of the world.
- *Put-me-here* travel is a basic method that simply takes the user to some specified position. This can be somewhat natural, like telling a cab driver your destination and arriving some time later, or this method can be totally unnatural such as selecting a destination from a menu and popping there instantaneously.
- *Orbital-viewing* is the least natural form of travel. In this method, the world (which often consists of just a model-sized collection of objects) seems to orbit about users depending on which direction they look. When users look left, the object orbits to their left, allowing them to see the right side. Looking up causes the object to orbit above them, showing the bottom side.

FIGURE 1-22

Astronaut Rick Mastracchio practices shuttle mission tasks in virtual reality.

Image courtesy of NASA

Some of the above methods of travel aid users in their movement through the virtual world by constraining where they can go. This constrained travel is one of many ways in which a virtual world can be designed to help users find their way around. Other wayfinding aids include the provision of maps, paths in the world to follow, obvious landmarks by which to site one, and instruments such as virtual compasses, among others.

CONCLUSION

This chapter covered the history, background, and terminology associated with VR technology and applications. The following chapter will address issues related to applying virtual reality to a problem or for some other purpose. The chapter will discuss basic issues related to the application of virtual reality, how the application examples in this book were chosen, trends in virtual reality applications, background on how the applications are related to each other, and commonalities and differences to watch for as you read the application descriptions. We present a taxonomy of virtual reality applications and explain the visualizations of the application database that is on the companion website of this book.