EXECUTIVE SUMMARY

This report reviews current virtual environment (VE) interface technology, that is, the technology that allows a user to interact with a computer-generated synthetic environment. The goal for developing VE systems is to provide a user with multimodal, highly natural forms of computer interaction. Thus, the interface technology plays a critical role. By looking at over ninety available commercial products and nearly sixty ongoing R&D efforts, this report builds a picture of current interface technology capabilities and discusses how these may change in the next few years.

Why has so much recent interest focused on VE systems? Quite simply, the potential of these systems is enormous. First of all, they offer a more intuitive metaphor for human-computer interaction. The user can exploit his existing cognitive and motor skills for interacting with the world in a range of sensory modalities and, in many instances, the experience he gains in the VE is directly transferable to the real world. Also, VE technology opens up new application areas that, hitherto, have been too expensive, too dangerous, or simply impractical. While current examples of VE applications range from surgical training systems to futuristic adventure rides, the full scope of possible applications for VE systems, and their potential benefits, is still to be determined.

The technologies that are discussed are those relating to visual, auditory, tracking, primary user input (that is, glove-based, exoskeleton, joystick, trackball, 3-D mouse, and pen-based input), haptic, full-body motion, and olfactory interfaces. The role of visual interfaces is obvious and needs no discussion except to point out that humans are strongly oriented to their visual sense, even to the extent of giving precedence to the visual system if there are conflicting inputs from different sensory modalities. While tracking is a type of interface that is largely transparent to the user, it is critical in keeping the VE system informed about user movements so that sensory inputs can be correlated to the user's position. Auditory interfaces can play a key role in providing informational inputs to the user, increasing the realism of a simulated environment, and promoting a user's sense of presence in a VE. In addition, they are used in sensory substitution where, for example, a tone is sounded to indicate when a user comes "in contact" with a virtual object and so substitute for the sense of touch. The term primary user input interfaces is used here to refer to those means whereby the user provides direct input into the VE system, for example, commands that control the operation of the system. Haptic interfaces provide the tactile and kinesthetic feedback arising from user contact with objects in the environment. Full-body motion interfaces fall into two categories. Active self-motion interfaces allow a user to move freely

through an environment, for example, walking over various types of surfaces or climbing stairs as necessary. Passive motion interfaces reflect the use of some type of vehicle to move a user through the environment. The final interface technology to be discussed is that of olfaction, where odors are used to provide the user with additional sensory cues about his environment.

At the present time, visual, tracking, and primary user input interfaces are the ones best suited for practical VE applications. In each of these cases, there is a solid basis of commercial products for potential users to choose from. Auditory interface technology is on the verge of becoming ready for use in practical applications. Indeed, increased use of auditory interfaces is the major change anticipated in VE interfaces in the next couple of years. Haptic interface technology still is largely in the research domain. Although various haptic feedback devices have been developed and a few have been used in prototype applications, the only practical use of haptic interfaces that is expected to occur within the next two to three years is with devices that are purpose-built for highly specialized applications. Widescale use of this technology is unlikely within the next five years. With respect to full-body motion interfaces, there are several entertainment systems that support limited types of highly specialized movement. Support for more general types of active user movement is exclusively a research topic with a variety of different approaches being investigated. The next few years likely will see continuing work of this type, perhaps with some prototype applications being developed. Active motion interfaces are not expected to become suitable for general use within the next five to seven years. Current work on interfaces for passive motion is focusing on a new breed of motion chairs, which will probably become widely used by the entertainment market in the near future. Olfactory interface technology is the least mature of all the technologies discussed here and unlikely to see practical usage within the three to five year timeframe.

All current VE interface technologies suffer from some limitations, even the more mature visual, tracking, and primary user input technologies. In no instance does the interface technology match human capabilities for the relevant sensory modality.

In the case of visual interfaces, head-mounted displays (HMDs) are the primary means of achieving an encompassing visual volume. HMDs suffer from several problems, with the most serious limitations being:

- Inadequate display update rates when responding to user head movements.
- Inability to provide both high resolution and a broad field of view.
- Weight that imposes an inertial burden and low levels of comfort that prevent prolonged use.

All these problems are well recognized and the first two are likely to be substantially reduced in the next few years with advances in liquid crystal diode (LCD) technologies. While smaller, lighter weight displays will help to reduce overall HMD weight, the necessity for bulky optics means that weight will continue to be a problem. A former problem, the expense of commercial HMDs, is becoming less serious as more low cost devices are becoming available, although these require the user to make some compromises in resolution and/or field of view.

So far passive glasses have not been widely used in VE applications, although new microelectronic fabrication techniques for creating polarizing filters at the pixel level may change this trend. Shutter glasses are quite widely used, usually with cathode ray tube (CRT) or projection displays. Here again, advances in LCD technology are likely to see an impact as LCD displays with faster switching time will help in reducing crosstalk problems. There is much research and development in the area of autostereoscopic displays and a small number of products is likely to come to market in the next two to three years. Retinal displays are a new topic of research and development. While they have the potential for providing a fully encompassing visual display without the weight and limited resolution and field of view of current HMDs, it will likely be some years before these become available for practical use.

Systems for tracking head, hand, and body movements are available and many have seen widespread use. Even so, low latency, high accuracy systems for tracking in noisy, unprepared environments do not exist. The most serious shortcoming of current technology is the following:

- Inherent limitations in some combination of accuracy, intrinsic latencies, working volume, susceptibility to interference of obscuration, and cost.

Again, these are well-recognized problems that are expected to be the focus of near-term research and progress, especially for magnetic trackers, is expected. The most significant improvements in tracking performance, however, are expected to come from the use of hybrid trackers where many of the limitations inherent in a particular technology can be overcome. Only limited research is being performed on wide-area trackers and this type of tracking interface is not expected to see widespread use any time soon.

Eye tracking also is a less mature type of tracking technology. The major problems appear to relate to accuracy and intolerance to user head movements. The increased use of multimodal interfaces (in both VE and non-VE applications) that can benefit from the ability to monitor the direction of the user's gaze, however, is opening up new potential markets that should encourage further development of this type of interface technology.

A number of 3-D sound processors that can be used in VEs are commercially available. These range in capability from systems for use with PCs, to high-end professional audio systems. However, a number of questions need to be answered and further research done before virtual audio can become a practical tool. Serious limitations are the following:

- Inability to represent sounds as being located in front of the user and to adjust sound spatialization to head movements.
- Inadequacies in acoustic signal generation.

Near term work is expected to focus on these areas, continuing to improve the realism and full-surround capabilities of the technology. Crucial support for this work will come from the development of improved algorithms, based on a more thorough understanding of how humans perceive sounds. As digital signal processing becomes less expensive, virtual audio is likely to become more widespread; it is expected to become a common component of VE systems within the next five years.

The development of glove-based devices for user input is an area of current growth. The set of available products do allow the use of natural hand gestures for certain, limited interactions with a VE but the primarily shortcoming remains:

- Limited joint resolution and poor discrimination between gestures.

While improvements in sensor technology might help reduce this problem, it is likely that advances in software-based gesture recognition will play a more important role. Gloves already are a fairly common VE input device but their use is expected to become more widespread as gesture recognition capabilities improve. There seems to be little ongoing research looking at the use of exoskeleton-based devices and these are not expected to be widely used, but limited to highly specialized applications.

A fairly diverse range of 3-D mouse-based, joystick, trackball, and pen-based input devices is available. These products represent mature technology and, while new products may appear over time, no major changes in this area are expected.

Tactile and force feedback interfaces for VEs have been able to exploit previous work in the areas of, respectively, sensor substitution devices for the disabled and teleoperation. Both represent active areas of research and development. In the case of tactile interfaces, researchers are investigating how to provide contact force, slip, texture, vibration, and thermal sensations. Products intended to simulate contact forces that occur when a user touches a virtual object and products that provide temperature feedback are already commercially available. The ability to support other types of tactile sensation is more problematic. In addition to shortcomings in tactile interface hardware, much work is still needed in developing the software models needed to drive the generation of tactile signals. The major limitations in the area of tactile feedback can be summarized as follows:

- Limitations in the ability to represent surface characteristics such as texture, local shape, and slip.
- Inability of devices to present a range of tactile sensations.
- Limitation of tactile feedback to small areas.
- Lack of models and algorithms for efficient generation of tactile signals.

As stated, this is an active area of research and much progress is expected over the next few years. Nevertheless, although several prototype applications are expected, tactile interfaces are unlikely to see common use within the next two to three years. Practical applications should start appearing shortly thereafter.

The majority of current force feedback devices can be distinguished as exoskeleton devices that deliver forces to the shoulder, arm, or hand; tool-based devices that deliver forces to the hand via a knob, joystick, or pen-like object held by the user; thimble-based devices that deliver forces to the user's fingertips; or robotic graphics systems that move real objects into place to provide natural forces to the user. Each type of device is limited in the type of interactions it can support. Consequently, although several devices are on the market, each provides very different capabilities and is suitable for different types of application. The serious limitations of force feedback interfaces are, in many respects, similar to those given for tactile interfaces:

- Inability to provide force feedback for a variety of different VE interactions.
- Limitation of force feedback to a restricted number of joints.
- Intrusive nature of force feedback devices and their constraints of user movement.
- Lack of common models and algorithms for efficient generation of kinesthetic signals.

This too is an active area of research where technology advances can be expected to occur in the next five years. It is likely, however, that initial advances will be application-specific, largely in the area of medical applications where there is much interest in supporting the simulation of surgical procedures. Only a couple of the current devices have seen any practical use and more widespread use is not foreseen in the next few years.

A number of approaches and devices have been developed to facilitate a user "moving" through a VE. The simplest, and most common of these, is for the user to point in the desired direction and for the visual scenes to be adjusted accordingly. A number of entertainment systems provide highly specialized interface devices allowing, for example, the user to simulate hang gliding or sledding. Unfortunately, there has been little progress in providing more general interfaces that allow a user to simply walk or run through a VE. Technology that can support a user moving through a large area or across a surface with varying characteristics has only recently begun to be investigated. A number of diverse designs for interface systems have been proposed and a few prototypes built, using both mechanical and non-mechanical approaches. While such systems may see use as advanced prototypes, none are expected to come into common practical use within the next three to five years. The potentially large entertainment market also has fostered the development of passive motion interfaces. In the last year, several motion chairs have been developed that employ techniques ranging from inflatable chair cushions to motion bases in order to provide the user with a sense of motion. These devices may become widely used for a diverse range of low-cost simulators.

Attention only recently has turned to providing olfactory cues for VEs. There are a few commercial systems available, but none of these is capable of full control of the user's breathing space. Some prototype systems that do provide such control, at least one of which is intended to be portable, are being developed. Nevertheless, the demand for olfactory interfaces is relatively small and this technology is expected to mature slowly and not become practically available in the near future.

In addition to further research and development on actual interface hardware and software, all the areas of interface technology discussed in this report will benefit from a better understanding of the role of sensory cues and human perceptual issues. This improved understanding not only is required to know how sensory cues can be delivered or simulated, but when and how they should be used. This is not to say that full fidelity of sensory cues is the ultimate goal. Even if achievable, high levels of fidelity would be expensive and not always desirable. What is needed is to determine the fidelity required for specific applications and how best to satisfy those requirements.