# Chapter 15

# Virtual Reality

Suppose a surgeon needs to plan a brain operation to remove a tumor. 3D images of the patient's skull and brain are available from the diagnosis. Using *virtual reality*, the surgeon can practice with the 3D data model rather than the real object. Different entry routes can be examined and different physical operations tried in order to choose a best intervention for the real patient. Moreover, it is possible to correspond a generic "brain atlas" with the 3D image data so the surgeon can see it in overlay and evaluate the consequences of different surgical options, avoiding damage to critical brain structures. Virtual reality (VR) systems have been developed to facilitate such virtual surgery. Figure 15.1 shows a rendered image of a brain model that could be used in a VR system.

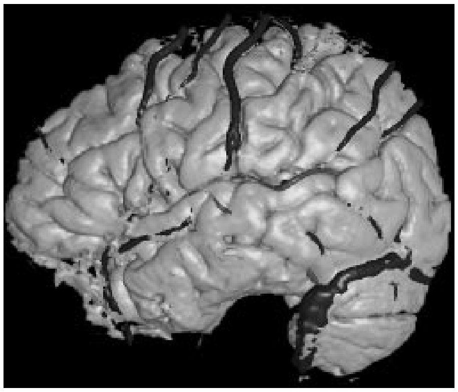


Figure 15.1: Artificial rendered image of a 3D mesh constructed from real patient MRI data (courtesy of the University of Washington Human Brain Project).

Virtual reality is a new field and often considered to be a subfield of computer graphics, because computer generated displays are an important component of a VR system. VR applications are important for study in their own right and VR systems also relate to the purpose of this text in multiple ways; (a) real images and image processing are often

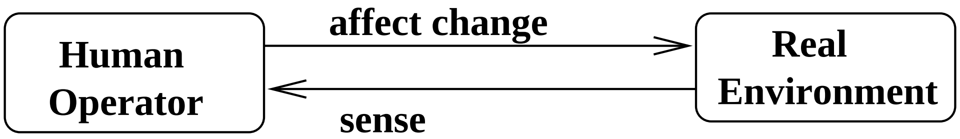


Figure 15.2: Human operator working in a real (natural) environment.

needed, (b) quality stereo imagery must be produced to *immerse* the human user in the virtual environment, (c) common mathematical models are used to correspond 3D points in various real and model spaces, and (d) machine vision is sometimes used to sense the position of the user or other real objects. VR has grown from roots in engineering of simulators, particularly flight simulators, teleoperation, and computer games.

## 15.1 Features of Virtual Reality Systems

We identify the important features of VR systems or VEs (virtual environments), after which we cite several exciting applications.

- A human user operates with a **model** of reality and simulates many of the operations that are possible with real objects.
- The display methods are of high quality in resolution and speed, so that the user is *immersed* in the data and problem to approximate perception of the real thing.
- The user must be able to interact with and change the model environment in smoothly perceived real time.
- 3D visual feedback is of paramount importance. The VR system typically provides a way for the user to change her viewpoint of the environment or to rotate or move objects for better viewing. Although visual feedback is most important, some tactile, motion, force or auditory feedback should also be present, so the user can feel objects or hear them collide, etc.

Figure 15.2 illustrates the usual scenario of a human operating in a real environment, while Figure 15.3 illustrates the newer scenario of a human operating in a virtual environment.

## 15.2 Applications of VR

VR is usually asociated with exotic new applications enabled by new expensive hardware. However, VR existed to some degree in several commonly known older systems.

### Architectural Walkthrough

A user might interact with an architectural model house by taking a virtual walk through the house, perhaps looking out the virtual windows at a virtual landscape. If the house were

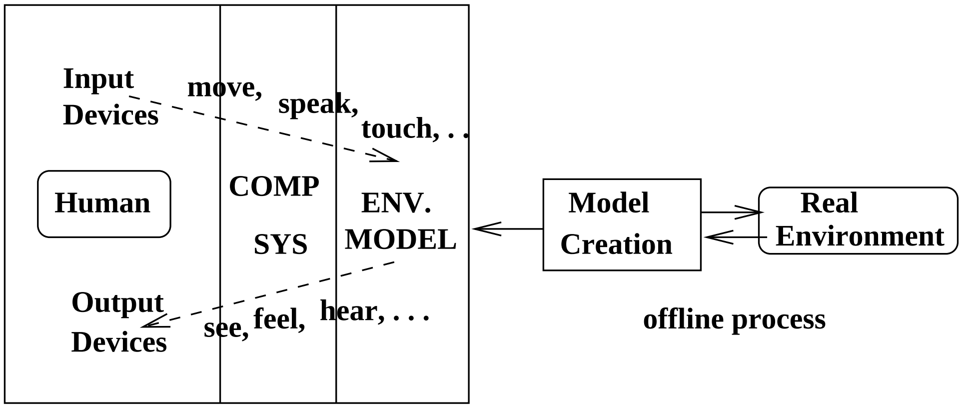


Figure 15.3: Human operator immersed in a virtual environment.

---

**Exercise 1** Books and Movies

- People can certainly become "immersed" in a plot when reading a book. Which of the above four features of VR systems are present and which are not? (Note that there are some books that allow the reader to select from different continuations of the plot.)

- People can also become immersed in a movie, especially when a wide screen or 3D display techniques are used. Which of the above four features of VR systems are present and which are not?

- Do you know of or play any video games that have all four features given above? Explain.

---

Jefferson's Monticello, the user could view Jefferson's collection of artifacts, his unique bed and cannonball clock. Many such historical places and *virtual museums* are currently being digitized. In simple cases, users can explore these virtual environments using the WWW and an ordinary flat display. Users might not be allowed to modify a model of Monticello, but they might be able to pick up and examine artifacts inside of it. If the user is planning to build a real house, then he should be able to make changes to the architectural model, perhaps experimenting with different wall coverings or furniture placement.

Related to the "walkthrough" is the "flyby": one can, say, view the Grand Canyon as if flying over it. In the simple case, the view cannot be determined by the user who is a mere passenger in the plane; in the more complex case, the user is the pilot who can dynamically change course and view the scenery in many different ways.

#### Flight Simulation

With a flight simulator, a user can practice control of an aircraft navigating various terrain and in takeoffs and landings at various airports. Thirty year-old stories tell of the increased pulse and perspiration of humans exiting flight trainers, attesting to the depth of their immersion in the virtual environment.

#### Interactive Segmentation of Anatomical Structure

A VR system can help medical personnel identify and extract models of anatomy from 3D sensed data; in other words, support interactive segmentation. Suppose, for example, that 3D MRI data is displayed to a user via stereo images. By interacting with the data, the user could mark a sequence of points at the center of a blood vessel, or on the boundary of a chamber of the heart. Stereo display devices and 3D input devices needed for this application are discussed below.

Today, there are many more applications of VR systems including pain management, phobia treatment, low-vision aids, driving simulators, scientific visualizations, and virtual classrooms. Figure 15.4 illustrates a current VR project at the University of Washington: the Dynamic Virtual Playground. The virtual playground is a prototype system designed to investigate multiple simultaneous collaborations in a virtual setting. It could be used to simulate a school lab where each group of students is working on a different project.

## 15.3 Augmented Reality (AR)

A contractor remodeling an existing site needs to know the locations of existing water and gas pipes and electrical conduit. Such data could be available from existing maps, blueprints, or even CAD files. The following scenario is an example of *augmented reality* (AR), also called *mixed reality*. The contractor wears a head-mounted display (HMD) that overlays computer graphics on the real scene that he is viewing. When he looks at the ground, he sees blue lines where water pipes are buried and when he looks at a wall he sees red lines for electrical conduit and blue lines for water pipe. In some sense, the AR gives the contractor superman's capability of seeing through walls!

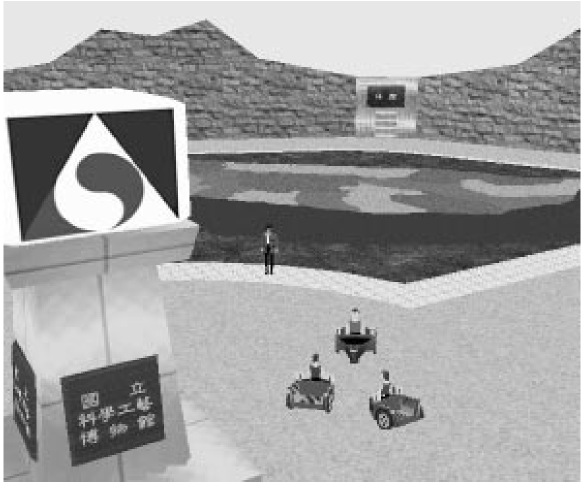


Figure 15.4: The Dynamic Virtual Playground is an experimental environment in which users can collaborate (courtesy of the University of Washington HIT Lab).

Creation of such an AR system requires the following.

- 3D models of the objects needed to augment the real view.
- Correspondence of the user's real workspace to the 3D model data via calibration.
- Tracking of the user's pose to determine the user's viewpoint within the real workspace.
- Real time display combining the real images and the computer graphics generated from the models.
- The response time to head movements and the registration accuracy between image and graphics, which are critical to the effectiveness of the system.

An augmented reality environment is sketched in Figure 15.5 and should be compared to the other disgrams of this chapter. There are many applications of augmented reality: here are a few.

- Consider AR-assisted surgery. The surgeon operating on a real patient views CAT scan data including a path plan superimposed on a live image of the patient. (The path plan might have been obtained using a VR system as above.)
- In PC board inspection, a human inspector compares a new PC board with a CAD model and verifies the presence of all required components and leads. The board is placed accurately in a jig so that its image from a camera can be accurately registered with the CAD model and presented to the inspector via a large computer display.
- The driver of an auto views a display that shows geographic features ahead. Projectors in the dashboard project the names of buildings and streets on the inside of the windshield.

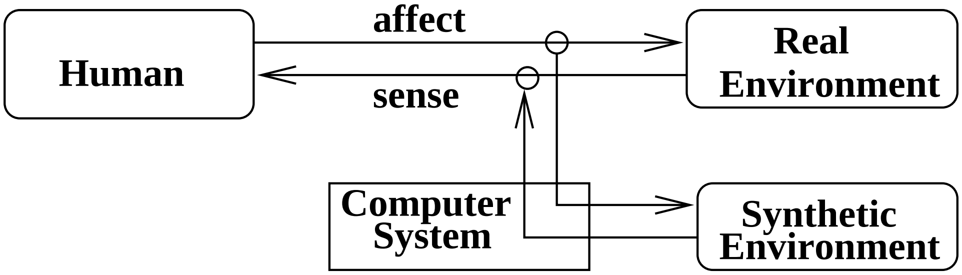


Figure 15.5: Human operator working in an augmented reality environment.

- Several people at a meeting want to talk about a computer model they are jointly creating. They want to be able to look at the model, point to and discuss its features, and still see one another and their surroundings. This scenario can be extended to teleconferencing where some of the people are located at a remote location. They may want to see not only the common computer model, but the other participants.

Figure 15.6 illustrates a potential teleconferencing application. The user of this system has placed two cards on his desk. Each card has a white background, a black square region, and a pattern inside the black region. Computer vision techniques are used to find the cards and statistical pattern recognition techniques are used to recognize the patterns within them, thus identifying the meaning of each card. Using augmented reality techniques, an image of the remote collaborator is placed on one of the cards, while an image of a graphics model to be discussed is placed on the second.

Figure 15.7 illustrates the case of two people in the same room using see-through goggles to look at a set of web pages they are both discussing.

## 15.4 Teleoperation

*Teleoperation* is an established engineering discipline which has provided much to virtual reality; in particular, the sensors and effectors needed to couple a human agent to the environment. The sketch in Figure 15.8 should be compared to the others of this chapter for similarities and differences. Using teleoperation, a human performs real work in a real environment that may be remotely located. A robot or robot-like machine carries out operations in the real environment according to the human control. Example successful application are as follows.

1. From a computer at Pathfinder mission control, a human operator gives a command to a navigating robot on Planet Mars to advance 10 cm and take a soil sample. The human gets feedback from images taken from cameras on the nearby lander that delivered the robot to Mars. Transfer of images and commands takes about 11 minutes due to the long length of the link.

Figure 15.6: An application of augmented reality techniques to teleconferencing (courtesy of the University of Washington HIT Lab).

2. Communicating via radio frequency, a human operator controls a remote robot that is vacuuming up radioactive waste in a dangerous area of a nuclear power plant after a minor accident. The human wears a head-mounted display which shows the contaminated area as the robot cameras see it and uses a simulated shaft to simulate the real vacuum shaft. Sensors in the simulated shaft sense the position and motion of the shaft yielding control signals to control the motion of the remote real vacuum.

3. A surgeon performs remote surgery by sewing a synthetic object similar to a football. Sensors record the delicate sewing motion parameters which are then transmitted to a remote robot that sews up an incision in a real live dog. This experiment has alrady been performed and may be a step in delivering the skills of a surgeon to otherwise unreachable places.

4. (Future scenario.) A patient goes to a hospital to have artieries cleaned of plaque (arteriosclerosis). The person is placed in an MRI machine which enables real-time 3D viewing of the body. Microrobots are injected into the blood stream and allowed to disperse throughout the body. The attending doctor, aided by maps from previous diagnoses, indicates via a 3D input device what regions of the vasculature need cleaning. The MRI device then operates in alternating modes, one imaging as before, and the other stimulating the microrobots to perform cleaning actions in the designated areas.

Figure 15.9 shows an example of a real, working telerobot system from Kyushu Electric Power Company. The robots are used to safely repair high-voltage electrical power lines.

Before moving on to discuss the devices and mathematical models needed to implement virtual environment technology, we give a humorous and instructive anecdote. A teleoper-

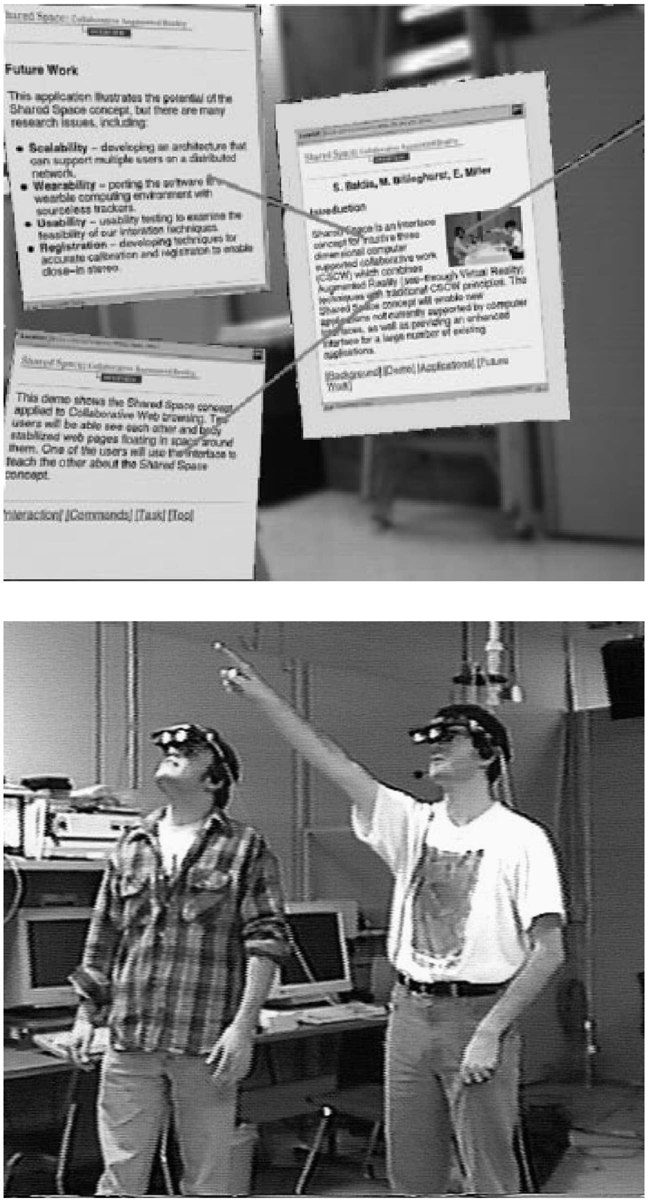


Figure 15.7: Two people wearing see-through, augmented-reality goggles (below). They can see the real world, and they can also see computer-generated displays. They are both looking at the web pages shown above, which are perceived to be floating in space. (Courtesy of the University of Washington HIT Lab.)

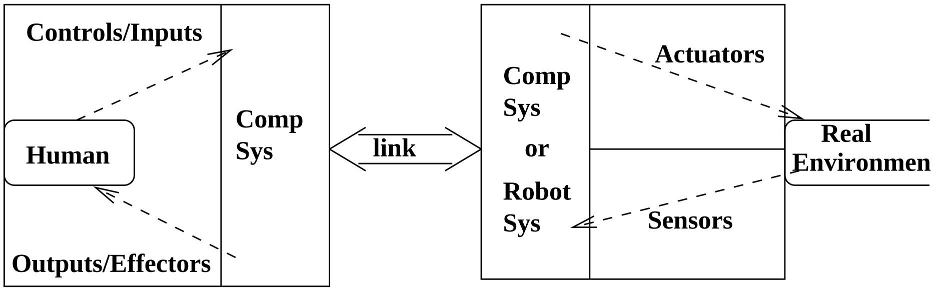


Figure 15.8: Human operator using teleoperation.

Figure 15.9: The "Hot-Line Telerobot System" from Kyushu ELectric Power Co. A teleoperated robot repairs a high-voltage power line (left). The operator's interface to the system (right). (Courtesy of Blake Hannaford with permission of MIT Press. Reprinted from K. Goldberg, *The Robot in the Garden*, Cambridge, MA: The MIT Press, 2000.)

ated power shovel was devised so that a remote operator could move soil, coal, etc. The operator wore a *data glove* which could sense the position of his hand and fingers, which the system then translated into control signals for the power shovel. Two cameras on the power shovel provided left and right images for the operator's HMD. The operator would use his hand to grasp in a pile of sawdust on a table, and the remote power shovel would mimic these motions by grasping in a pile of actual coal. Once the human operator had an itchy nose and reached to scratch it with the data glove hand! Moving the controlling hand toward the nose caused the power shovel effector to move toward the cameras in the real environment. When the resulting images were transmitted back to the operator via the HMD, he perceived that he would be struck in the face by a massive shovel! This was only a virtual blow and the operator was not physically harmed; however, he was psychologically upset enough to need a break from the work. ( Had the system not been carefully designed, the cameras on the real power shovel might have actually been damaged.)

1 DEFINITION *Using* **teleoperation**, *a human operates a real device in a remote real environment: the feedback from the remote environment and the controls of the device provide some illusion that the operator is present in the real environment.*

2 DEFINITION **Virtual reality** *is a synthetic reality provided by a computer system to a human user via rich models of reality and immersive input/output devices: the human operator has the illusion of working with real objects when none are actually present. The illusory environment which the operator perceives and changes is called a virtual environment (VE).*

3 DEFINITION *An* **augmented reality** *or* **mixed reality** *is created by combination of a real environment and a virtual environment: synthetic outputs from a computer system are combined with sensed data from a real environment to augment the human's perception of reality.*

4 DEFINITION *A* **synthetic environment (SE)** *is an environment provided to a human operator by a computer system and immersive I/O devices via teleoperation, augmented reality, or virtual reality. (Sometimes, the term virtual environment (VE), is also used to cover all of these cases.)*

## 15.5 Virtual Reality Devices

Several devices are commonly used to couple a human operator to a synthetic environment. Figure 15.10 guides our discussion: it shows a previously discussed AR application, but the devices are used in all three types of SE applications. A contractor walks about in a building to be remodeled and marks the walls where water pipes and electrical conduit are present. The contractor sees real walls via a see-through HMD containing optics to overlay computer generated images showing the pipes and conduits.

The contractor (a) interacts with a real environment (b) augmented by a computer system (c) which adds otherwise invisible features to the images seen (d,e,f) and uses speech communication (g,h) with the contractor, who is free to walk around and mark the walls (i). As the human moves about, the image of the real world is seen through a beam-splitter (d). A pose sensor (e) transmits the orientation and location of the head to the computer,

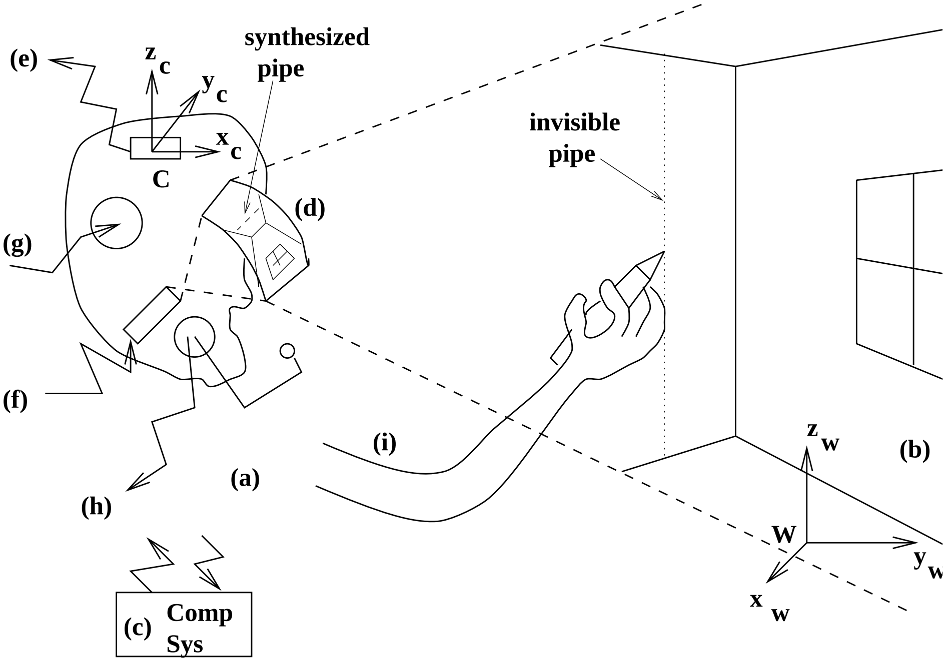
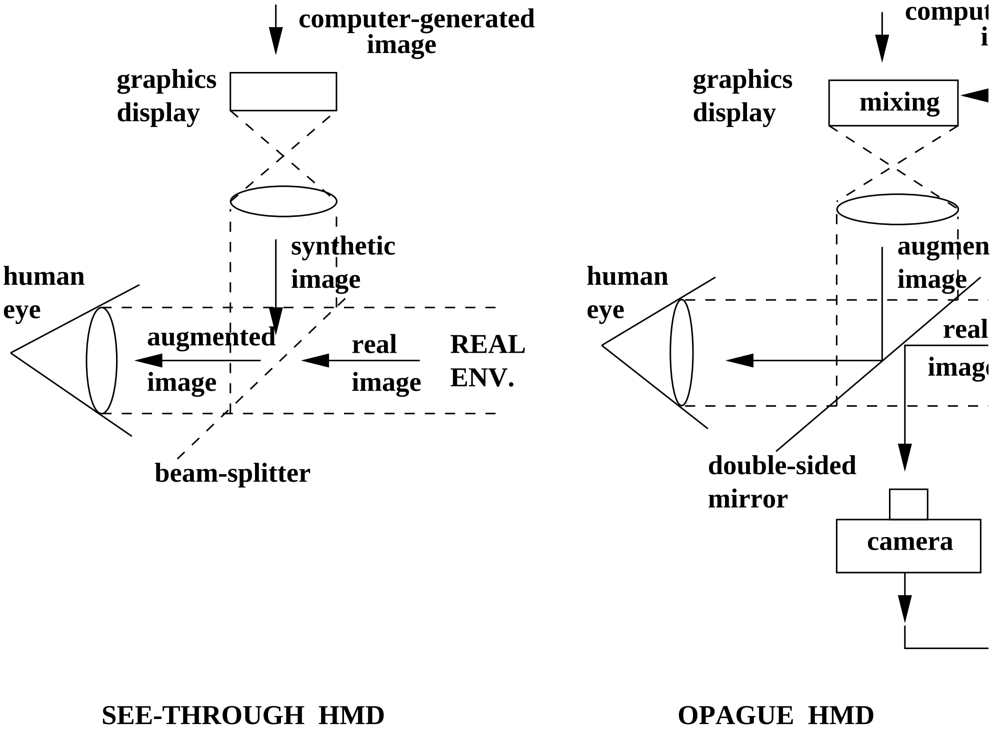


Figure 15.10: Devices used by a human operator in an AR application. The operator marks the location of pipes in a wall: location of the invisible pipes is known from a CAD model of the house.



SEE-THROUGH HMD

OPAGUE HMD

Figure 15.11: Schematic of head-mounted display (HMD): (left) see-through and (right) opaque.

which then uses these parameters to generate a computer image from a 3D CAD model of the pipes and conduits. The generated image is projected onto the inside of the one-side mirror through which the operator is looking (f) to augment the view of the real world.

### The Head-mounted Display

A schematic of the *see-through* HMD is given at the left of Figure 15.11. A beam-splitter allows light from the real-world to pass through from the outside, but acts as a mirror on the inside in reflecting the light from a computer-generated display, resulting in the operator seeing the augmented image. Note that the graphics display is a miniature display with small optics residing in the head-mounted display or helmut. An alternative design, the opaque HMD, is shown at the right of Figure 15.11. Note that all elements reside in the HMD and move with it – the mirror, the camera, and the graphics display. Here the synthetic image is a mixture of a digital image from the camera and a computer generated image. The see-through design can provide higher resolution images, while the opaque design provides more options for controlling what the user sees.

**Exercise 2** On registration accuracy for AR systems

Refer to Figure 15.10. a) Suppose that the pose sensor makes an error of two degrees in pan angle when reporting the pose of the operator's head. What would be the error (in cm) in the real world between where the operator would mark a vertical pipe versus its real location? b) Suppose that the visual display has an FOV of 120 degrees mapped across 500 pixels. (If needed, you may assume a focal length of 2.0 cm for the imaging system. Also, you may assume the standoff from the wall being observed is 3 meters.) What would be the horizontal distance, in pixels, between the projection of the pipe in the augmented image and the true location of the pipe in the image?

**Exercise 3** On registration accuracy for AR systems

This problem relates to the previous one and uses the same devices. The problem is inspection of automobile instrument panels via human comparison of a real instrument panel with a CAD model using an AR system. Unlike the previous problem, all CAD features should be visible to the human operator in the real image. The operator is to check the existence and correct functioning of the instruments, such as odometer, oil pressure guage, radio, etc. The AR computer system also drives the test equipment and gives input to the operator prompting the operator on what to look for. Assume an FOV of 60 degrees and a standoff of about 60 cm. As in the problem above, suppose the pose sensor has an error of two degrees in the pan angle. (a) If the augmented image has a red circle where a radio knob should be, what is the horizontal error in registration caused by the error in pan angle? (b) Assuming misregistration error is bad, describe a method to automatically reduce it using the computer system. Would either type of HMD work as well for your method? (c) Assuming that the operator could do better inspections by being able to control small misregistrations, describe a method to automatically produce small misregistrations using the computer system. Would either type of HMD work as well for your method? (d) Is an AR system needed for this problem or could it be solved with a completely automatic system? Explain.

**Exercise 4** Multiple operator AR

Suppose a team of several surgeons will do surgery on a patient. Is it possible for each to wear an HMD to view the surgery plans and anatomical structures overlaid on their view of the real patient? Explain how it might be done or why it can't be done.

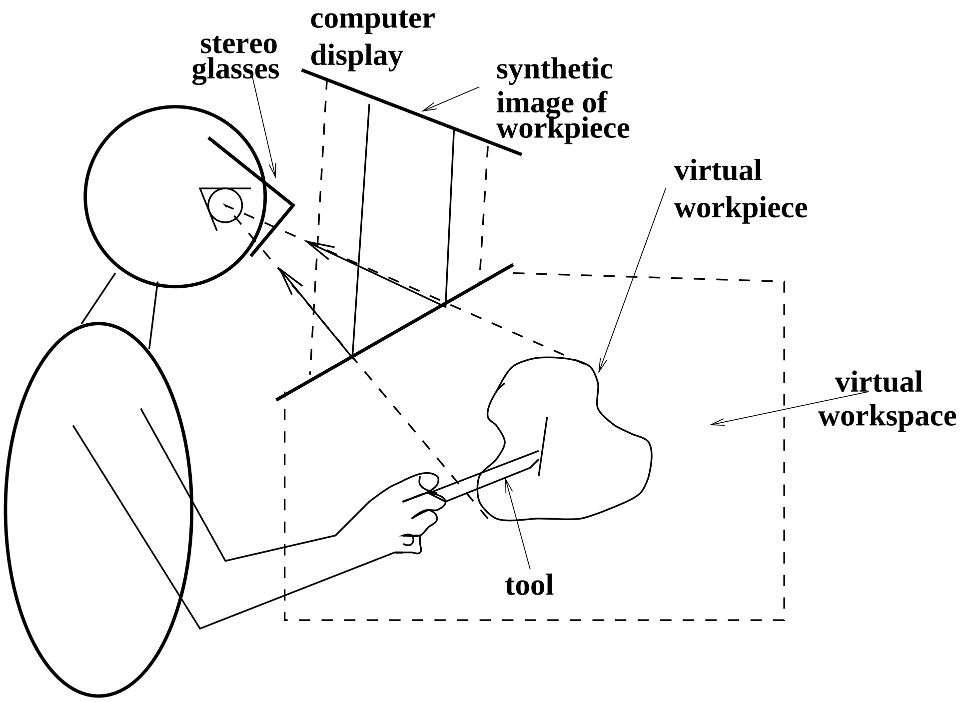


Figure 15.12: Schematic of workbench for dextrous virtual handwork.

### Dextrous Virtual Work

A VR system supporting dextrous handwork on models is sketched in Figure 15.12. The human operator views the synthetic workspace via a stereo display projected onto a mirror from above. This allows normal freedom of hand movements in a real 3D workspace below the mirror. The operator manipulates a tool below the mirror; the pose of the tool is carefully tracked and its image is projected back onto the mirror so that the operator receives visual feedback. The figure shows an operator practicing making an incision on a virtual organ. Different technologies available for the 3D tool are given below. Clearly, accuracy of 3D pose and speed of sensing and update of the display are critical to the system. Besides obvious applications for the practice or plannning of surgery, such a system might be useful for an artist in digitally sculpting a 3D model.

### Stereoscopic Display Devices

Stereo vision is perhaps the most important visual cue for sensing 3D objects that are within 10 meters and is the predominant means of feedback from VR systems. There are two common designs for stereo displays. With opaque HMDs, separate images can be presented to the two eyes; the left and right images are synthesized with the proper disparity from the object model using the mathematical model of Chapter 12. Opaque HMDs can produce a wide FOV on an infinite virtual world and hence high degree of presence/immersion; however, there are many design constraints in creating devices that will comfortably fit different human operators. Note that multiple users can be immersed in the same virtual world: each user has an HMD with individual pose sensor, providing an individual immersive display of the VE from the user's viewpoint. An alterative is to use an ordinary graphics display as shown in Figure 15.12 which displays the left and right images interleaved over time. The operator uses shuttered glasses that are syncronized in time with the display, so that the left eye sees only the odd frames and the right eye sees only the even frames, or vice versa. The perceptual system is able to fuse this time-varying input to perceive 3D. This design permits an inexpensive system that is easy to use. However, the degree of immersion is limited because all imagery is limited to the display screen (called *fishtank virtual reality* for obvious reasons): the view of the virtual world is limited by being clipped to the display screen (*fishtank*). Moreover, since there is only one physical display synched to one user, other users viewing the display do not have their own viewpoint and do not benefit from their own motion. The fishtank effect can be reduced by surrounding the user with displays, thus creating a *CAVE*, but the full head-coupled stereoscopic effect is still unavailable to multiple users.

## 15.6 Summary of Sensing Devices for VR

### Visual

Visual output from a VR system, as described above, typically consists of stereoscopic displays to the user, either via the HMD or standard computer display viewed through shuttered glasses. This output is coupled to the user via sensors that sense user pose in the VE so that the display portrays the appropriate viewpoint.

Visual input to a VR system, if present, usually consists of tracking the users eye[s], head, or body parts and providing pose as inputs to the activity being modeled. Eyetrackers can provide gaze direction. Devices can be incorporated into HMDs or can be totally separate from the operator. Tracking head pose using HMDs is possible using a camera that tracks special feature points on the HMD. Recent work in VEs has demonstrated the capability of tracking the hands, head, and feet, or limbs using the best views from multiple cameras observing the human. One advantage of visual input devices is that they need not be another constraining wearable device; however, although the user's movement may be unconstrained by the device, the devices still have limited work volumes. Some eye trackers depend on special IR illumination and some body trackers depend on controlled background colors.

### Pose

3D position sensors sense the position and orientation of human body parts or tools that are held. 6 degree-of-freedom sensors include the Polyhemus sensor commonly used on HMDs, joysticks, and newer devices such as the *bat* described by Green and Halliday (1996). x-y-z position sensors include the sparking styllus and various mechanical devices. It is also possible, although not common, for computers to output pose to a human by positioning mechanical joints that are attached to the human. This is related to force feedback, which is described below.

### Auditory

Speech input to a computer has been available for over 15 years and is now appearing in many interfaces, including telephone systems and home PC applications. Speech input has the advantage of being natural, eg not requiring special learning, and may be necessary when an operator's eyes and hands are involved in other tasks and unavailable. Similarly, speech output is a convenient communication channel independent of the display.

Auditory output might enhance any interface – for example, a metalic banging sound confirms that a file has been deleted when a folder icon has been dragged to the trash can icon. It also enhances the degree of immersion in a VE: the operator of a virtual vehicle can hear the sound of the engine and the wheels squealing to a stop. Or, a teleoperator can sense the amount of radioactivity present via the degree of frenzy encoded in music being played.

5 Definition **Sonification** *is the encoding into sound of data or control information that is not naturally sound.*

### Haptic

Humans connect to the real world via their via touch, force and motion. Sense of touch is due to nerves in the skin that can sense temperature, firmness, and smoothness of surfaces. Nerves in the limbs and muscles can sense limb pose and muscle tension/length and their changes. Nerves in the vestibular system can sense body motion.

6 Definition *The human* **haptic sense** *involves the sense of touch (tactile sense) and the senses of body position, forces, or motion (kinesthethic sense).*

A variety of electro-mechanical devices are available to provide input and output of forces.

### Motion

Human perception of motion is affected by the integration of many sensing systems. Visual displays, as in the movies or simulators, can provide enough stimulation to actually induce motion sickness. Kinesthetic stimulation via treadmills, robotic linkages, or platforms or centrifuges that actually move the operator is employed in various VE systems to increase the degree of immersion above that provided by the visual display. Alternatively, the vestibular system might be artificially stimulated using a stream of cool air or water. Computer vision can play a strong role in motion sensing. Ideally, a human operator would move freely in a real environment and analysis of images from several tracking cameras would provide all the needed interpretation of motion and perhaps map it onto a computer model of the human body. Commercial systems are available for tracking the moving human body; typically, they depend upon placement of specially reflective targets at special locations on the body in order to simplify image segmentation and feature extraction. Such systems are used in various sports studies and in prescribing orthopedic devices. Research systems have demonstrated tracking of the human head, hands, and feet without use of special targets on the body.

## 15.7 Rendering Simple 3D Models

In order to create virtual scenes, we need the tools for constructing object models and for producing images of them. The object models used can be complex mesh models as discussed in Chapters 13 and 14, or they can be simpler wire-frame models, which can be created using Computer-Aided Design (CAD) software packages, such as AUTOCAD. Figure 15.13 shows a wireframe model of a car created by a user interacting with a CAD package.

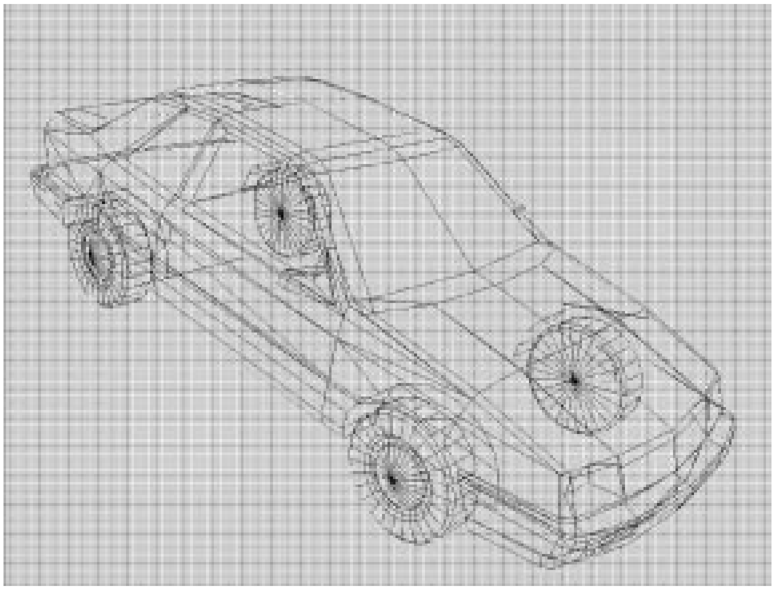


Figure 15.13: Wireframe model of a car (from the VRML Repository: http://www.web3d.org/vrml/oblib.htm.)

Once we have the 3D model, we would like to be able to display it from arbitrary viewpoints and with various different lightings.

7 Definition **Rendering** *is the process of creating images from models.*

Rendering can be thought of in two steps:

1. Determine which surfaces of the model are visible to the selected viewpoint.
2. Determine the corresponding pixel values of the created image.

Figure 15.14: Rendered image of the car alone.

Conceptually, step 1 is achieved by constructing a ray from the viewing point in the desired direction to the object. The first intersection of the ray with the object is the surface point that will be visible along that ray. This concept is called *ray tracing*, and there are many software algorithms to perform it. Today's computers use a hardware mechanism called a *z-buffer* to carry out this step rapidly.

Step 2 can be simple or very complex. In the simple case, the object has a particular color and is made of a particular material whose reflectance properties are known. The light comes from a point light source in a particular direction. Some amount of ambient light is also present. Mathematical models such as the Phong shading model given in Chapter 6 are then used to determine the pixel color corresponding to a small area of the object. Figure 15.14 shows a simple rendering of the wireframe car model. More realistic images can be created (at the cost of speed) by adding such factors as multiple light sources, areal light sources, shadows, transparent surfaces, and interreflections.

Figure 15.15 shows two rendered images using the car model and other models. While some of the objects in these images were rendered as described above, some of them have textured surfaces involving complex patterns that would be very slow to render. (See the building on the far left of the left image and the sidewalk next to it.) These surfaces have been *texture mapped* instead of being rendered. Texture mapping will be discussed in the next section.

Figure 15.15: Scenes using the car and other models.

## 15.8 Composing Real and Synthetic Imagery

Synthetic rendering can only go so far. Not only is it not sufficiently realistic, but attempts at realism make it very time-consuming. Instead, existing images of complex textures can be used to both improve realism and speed up the rendering. For this purpose, a *texture* can be an artificially generated pattern or a (portion of a) real image. In the process of rendering a surface, instead of coloring it with a single color value, we would like to 'paste' or 'paint' a given texture onto it. This leads to the process of *texture mapping*, where the final values of the visible pixels are selected from the pixels of a given texture.

8 Definition **Texture mapping** *is the process of painting a texture onto a smooth surface, thus creating a textured image of that surface.*

The surfaces of Figure 15.15 that were texture mapped were planar polygons, the easiest surfaces to texture-map. Texture mapping can also be done on more complex curved surfaces, such as painting a rough "peel" texture on an orange. When the object is free form and represented by a mesh, we can do this in a piecewise fashion. But for difficult objects, we need more sophisticated methods. A recently developed technique in computer graphics uses real images of the object to provide the needed textures. Figure 15.16 (a) shows a rough mesh model of the reconstructed dog (from Chapter 13), and Figure 15.16 (b) shows a texture-mapped image of the model. In this example, the texture comes from a real image of the dog taken from the same viewpoint at which the model is displayed. We really want to be able to display texture-mapped images from arbitrary viewpoints, as we did with rendering.

9 Definition **Image-based rendering** *is a technique that uses a set of real images of an object to produce an artificial image from an arbitary viewpoint.*

Image-based rendering can actually be accomplished without a geometric model of the object, by storing a very large number of images at different viewpoints and interpolating between them to produce arbitrary views. However, if we have a mesh model of the object and a small number of real images that cover the set of potential viewpoints, then the geometry plus the images can be used to produce very realistic renderings. The reconstruction system described in Chapter 13, which produced rough mesh models of objects from a set


(a) (b)

Figure 15.16: Rough mesh model of a dog and texture-mapped image.

of range images and associated color images, can also render the objects from various user-selected viewpoints, using a technique called *view-based texturing*. Figure 15.17 shows the basic principle behind this approach. On the left, a false-color rendition of the object can be manipulated by the user with a mouse to rotate it to a desired viewpoint. In the center, the three closest views to that viewpoint have been retrieved. On the right, a texture-mapped image of the dog in the desired orientation is produced. To create each nonbackground pixel of the image, a ray is sent from the image pixel to the 3D model and then from the model to the appropriate pixel of each of the three closest views. The colors from these three pixels are blended to produce a value for the selected pixel in the created image. Rather than counting all three pixels equally, the blending algorithm takes into account how similar a stored view is to the required view, the direction of the ray from the object model to the stored-view pixel, and the nearness of the stored-view pixel to the boundary of that view. It also uses a software variant of z-buffering by throwing out the contributions of pixels that are too far away to actually lie on the surface.

Full 3D mesh models of real objects take a long time to construct. View-based texturing does not actually need the full geometric model; instead it can work with only the original set of registered range/color images. Figure 15.18 illustrates this process for the dog model. Instead of a full mesh, a partial mesh is created for each of the sample views. These partial meshes can be used together with the color images of the object to produce renderings that are just as good as from a full mesh.

Furthermore, there are some real objects that are not suitable for producing solid models. When an object has thin parts, such as the sail of a boat or the leaf of a plant, the mesh model would have to be of extremely high resolution to capture the topology. Since it *is* possible to capture range/color data from a number of views, view-based texturing can still be used to produce very realistic images of the object from arbitrary orientations. Figure 15.19 illustrates this process on a basket of flowers.

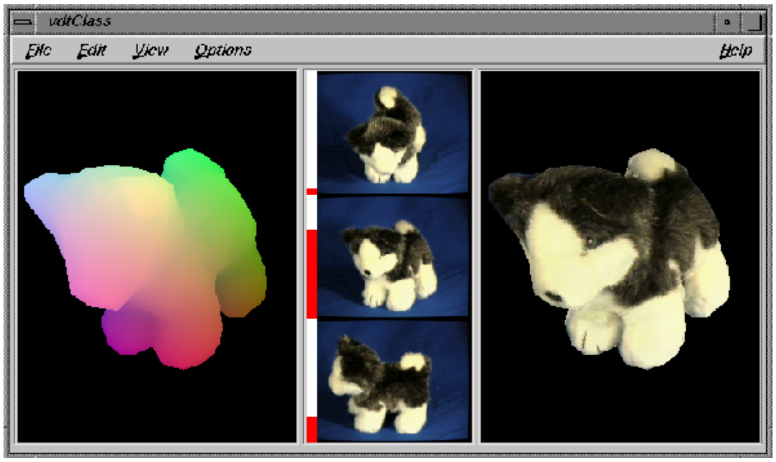


Figure 15.17: Range data of a dog model (left), three real color images from nearby viewpoints (center), and the rendered image using a weighted combination of pixels from these views (right).

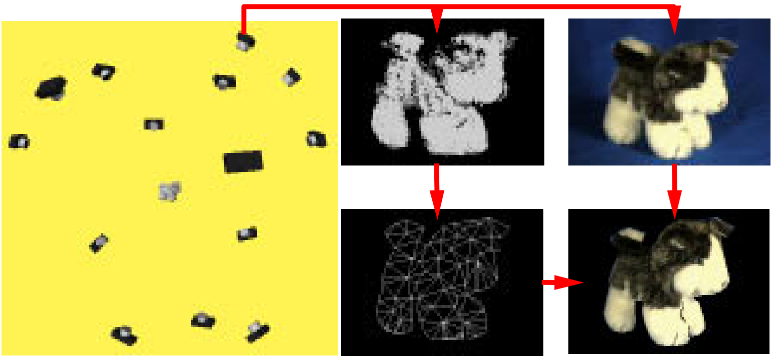


Figure 15.18: Registered range and color images from a small number of views of an object can be used to produce a high-quality rendered image, without ever constructing a full model of the object. Potential viewpoints (left), range data from one of the viewpoints (top center), color data from the same viewpoint (top right), mesh constructed from the range data (bottom center), and rendered image achieved by texture-mapping the color data onto the mesh (bottom right).

Figure 15.19: The same technique applied to an object for which construction of a full 3D model is nearly impossible, due to the thinness of parts of the object. Three different color images of the object (upper left), three images from a different, selected viewpoint constructed by mapping the pixels of the three original images onto the new viewpoint (lower left), and final rendered image, which is a weighted combination of the three constructed images (right).

---

**Exercise 5** Augmenting the image of a real cube

Calibrate a camera using a method from Chapter 13 and a calibration *jig* which is a cube or small box (use the 7 visible corners as control points). Design a 2D die face containing three dots to represent the number 3. For the top of the 3D cube, construct a mapping function $g$, as in Chapter 11, which maps a point $[x_m, y_m, z_m]$ from the top of the 3D cube to a 2D point $[x_t, y_t]$ within the square modeling the die face. This mapping should be linear and map the 4 corners of the top of the cube to the four corners of the die face. Generate and print the image containing the pixels from the real calibration scene everywhere except for the top of the cube, which should contain the pixels from the model of the die face. Repeat the above process using the same mapping function $g$ but using an arbitrary square scanned image, perhaps of your own face.

---

**Exercise 6** Synthesizing the image of a cube

Use a real camera matrix obtained as in the problem above, or use one from Chapter 13. (a) Create a synthetic image of a cube posed somewhere in the FOV of the camera. (b) Create an image as in part (a), except that one of the faces of the cube should be the face of the die as in the previous problem. (*c) Texture map two of the faces with two different photos of human faces.

## 15.9 HCI and Psychological Issues

Clearly, there are many opportunities to enhance the quality and bandwidth of the interface between man and machine using devices described in this chapter. In VE, the primary goal is to provide a quality immersive experience as intended. Quality interaction using sight, touch, and force is needed for virtual surgery, for example. Individual differences among humans create difficulties in engineering VR systems; for example, the different sizes and shapes of the head complicates the design of HMDs and vision system differences complicates the control of displays based on stereo fusion. Also, different people have slightly different perceptions of objectively identical stimuli for color, roughness, sound, etc.

VR systems can produce unintended effects. For example, motion sickness is usually unintended. Other unintended effects can be eyestrain, fatigue, or frustration; for example, all these can result from a stereo fusion subsystem that is not quite matched to the operator or to reality. Even worse, human operators of flight simulators have been known to suddenly become lost in their virtual environments! These issues provide many challenges to VR system designers.

## 15.10 References

An overview of the state of the art in virtual reality is provided in the 1995 report edited by Durlach and Mavor. A broad set of papers reporting recent work on various aspects of VR is included in the 1995 collection edited by Barfield and Furness. Both books provide extensive background, including description of devices, definitions, excellent examples and figures and categorized references to the literature.

Virtual reality has diverse applications in medicine. Discussion of its use in planning surgery and in interactive segmentation of 3D imagery is given in a paper by Posten and Serra (1996). Our treatment of dextrous virtual work drew heavily from the ideas in that paper. Use of VR in rehabilitation and in therapy treating various phobias is described in a set of papers edited by Strickland (1997). Creating object models and defining their behavior and motion is a difficult problem of current research effort. For a description of the problems and some methods for their solution, see the papers by Green and Halliday (1996) and Deering (1996). The immersive power of VR systems can not only help humans, but it can also make them sick or irritated, especially if stereo and motion imagery is not very carefully presented: see the article by Viire 1997. The book by Stuart (1996) contains many useful tables summarizing the properties of human sensory capabilities and the input and output device characteristics needed for VR systems. The discussion of view-based texturing comes from the work of Pulli (1997). A recent collection that includes some of the work discussed here is **Mixed Reality** by Ohta *et al.* *Mixed reality* is another term for *augmented reality*, the merging of real and virtual worlds.

1. W. Barfield and T. Furness III (Eds) (1995) **Virtual Environments and Advanced Interface Design**, Oxford University Press.

2. F. Biocca and F. Levy (1995) **Communication in the Age of VR**

3. M. Deering, it The Holosketch VR Sketching System, **Communications of the ACM**, Vol. 39, No. 5, May 1996, pp 54-61.

4. N. Durlach and A. Mavor (Eds) (1995), **Virtual Reality: Scientific and Technological Challenges**, National Research Council, National Academy Press, Washington, D.C.

5. M. Green and S. Halliday (1996), *A Geometric Modeling and Animation System for Virtual Reality*, **Communications of the ACM**, Vol. 39, No. 5, May 1996, pp 46-53.

6. Y. Ohta and H. Tamura (Eds) (1999), **Mixed Reality: Merging Real and Virtual Worlds**, Ohmsha, Ltd, Tokyo, Japan: also distributed by Springer-Verlag, New York.

7. T. Poston and L. Serra (1996), *Dextrous Virtual Work*, **Communications of the ACM**, Vol. 39, No. 5, May 1996, pp 37-45.

8. K. Pulli, M. Cohen, T. Duchamp, H. Hoppe, L. Shapiro, and W. Stuetzle, *View-based Rendering: Visualizing Real Objects from Scanned Range and Color Data*, **Proceedings of the 8th Eurographics Workshop on Rendering**, June, 1997.

9. D. Strickland (Ed) (1997) *Special Issue on VR and Health Care*, **Communications of the ACM**, Vol. 40, No. 8, Aug 1997.

10. R. Stuart (1996) **The Design of Virtual Environments**, McGraw-Hill, New York.

11. E. Viire (1997) *Health and Safety Issues for VR*, **Communications of the ACM**, Vol. 40, No. 8, Aug 1997, pp 40-41.