Virtual Reality

Edward Cheung
email: icec@polyu.edu.hk

29 July, 2003.

Contents

Four Key Elements of Virtual Reality Experience

- Virtual World
  - An imaginary space often manifested through a medium.
  - A description of a collection of objects in a space and the rules and the relationships governing those objects.
- Immersion
  - Immersion into an alternative reality or point of view (POV).
  - Sensation of being in an environment; mental or physical.
  - Mental immersion is the state of being deeply engaged; suspension of disbelief; involvement. The goal of most media creators.
    - The mimesis of a novel – a term indicating how real or at least consistent with itself a story world is. – Mental, third person POV.
    - Sense of presence.
  - Physical immersion is the state for bodily entering into a medium, synthetic stimulus of the body's senses via the use of technology; this does not imply all senses or that the entire body is immersed or engulfed.
  - In VR, the effect of entering the world begins with physical.

Four Key Elements of Virtual Reality Experience (cont.)

- Sensory Feedback
  - An essential ingredient of VR
  - Most case is the visual sense that received feedback
  - Haptic (touch) experience is possible
  - Need high speed computer as a mediating device
- Interactivity
  - Response to user actions
  - Interactive fiction; Zork from Infocom, Inc.

History of Virtual Reality

- The Ultimate Display – Ivan Sutherland, 1965
  - Describe a virtual world provided by the computer
- Head Mounted Display (HMD) – Ivan Sutherland 1968
- Users are presented with left and right views of a computer generated 3D scene. The user's head movements caused corresponding spatial changes in the images.
- Evans & Sutherland Corp. manufacture Image Generators for pilot training starting in 1968
  - http://www.es.com
- More flight simulators are available from Thales Training & Simulation
  - http://www.ttsl.co.uk

Applications of VEs

- Engineering
  - Design from conceptual stage
  - Modelling
  - Simulation
- Training
  - maintenance training of sophisticate equipment
  - Flight simulation
- Education
- Medical and Health Care
  - Anatomy studies
  - Operation Procedures
  - Treatment of Mental Illness
- Entertainment

Type of VEs

- Virtual Environment (VE)
  - An interactive computer generated "world" seen from a first-person point of view.
- Immersive VE
  - A real time 3D synthetic environment that appears to surround the user in space
- Augmented Reality (AR)
  - A type of VR in which synthetic stimuli are registered with and superimposed on real-world objects; often used to make information otherwise imperceptible to human senses perceptible.
  - The use of transparent glasses on which a computer displays data so that the viewer can view the data superimposed on real-world scenes.
- Collaborative Virtual Environment (CVE)
  - CVEs are shared virtual reality spaces where remote users can participate in a simulation or closely coupled tasks.

Some Terms in Virtual Reality

- Special effects on films are not VR if the actors are not aware of the virtual characters. These special effects are examples of computer animation.
- Cyberspace
  - A location that exits only in the minds of the participants, often as a result of technology that enables geographically distant people to interactively communicate.
- Avastar
  - A virtual object used to represent a participant or physical object in a virtual world. Usually for CVE.
  - Adapted from Hindi, meaning the earthly embodiment of a deity.

A Generic Virtual Reality System

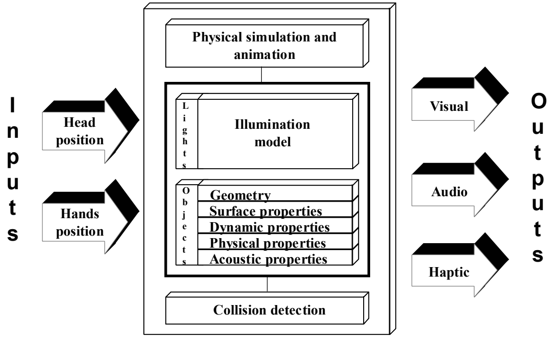


Figure: The inputs, processes and outputs in a generic VR system

Components of VR System

- Visual System and Display Hardware
- Audio System
- Tracking System
- Input Devices
  - Data gloves
  - Speech Input
  - Dedicated 3D devices; e.g. SpaceBall
  - Other devices, surgical apparatus, remote manipulators

Software

- Environment Model / Editor
  - A description of the scene to be rendered
  - Usually created in a commercial modelling package
  - May include behaviour specification
- Rendering or Display Software
  - Render 3D view of environment from viewer's viewpoint
  - May need collaboration with tracking devices
- Interactive Software
  - Navigation, selection, system control

Geometric Modelling - 2D Graphics

- The Cartesian coordinates provides an orthogonal presentation for points and area in 2D
- Graphics are composed by points, lines and curves
- Vector algebra can be used to manipulate graphics
- Display is 2-D
- Geometric Modelling or Raster Images
- Introduction of the z coordinates converted 2D data to 3 D and thus a virtual environment (VE) can be described
- In computer graphics and computer animation, the idea of a viewer or camera is used to view the object. Sometimes this is known as the Virtual Observer (VO)

Curves

- Bézier Curves
- Cubic polynomial curve segment
- Developed by Paul de Casteljau in 1959 and independently by Pierre Bézier in 1962 for Citroen and Renault as one of the ingredients in computer-aided geometric design (CAGD) system for automobile bodies shape design.
- A Bezier curve is derived from a sequence of control points that determine the curve's geometry. When 3 control points are used, a quadratic curve is formed between the first and last points. The middle control point influences the overall shape. When 4 control points are used, a cubic curve is formed.
- Slope continuity ensured smoothness and can be preserved over any number of segments by matching the starting slope of one segment with the trailing slope of the previous segment.
- Indirectly specifies the endpoint tangent vector by specifying two intermediate points that are not on the curve

Curves

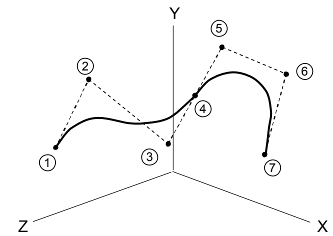
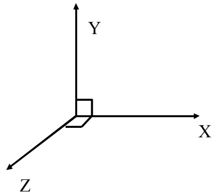


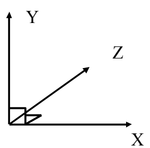
Figure: Long Bezier curves

Co-ordinate Systems

In computer graphics applications, vectors are used to describe scenes and to performing operations such as translation, rotation, transforming, etc.
The coordinate systems (right- and left-handed), serves as a reference point.



Right-Handed System
(Z comes out of the screen)



Left-Handed System
(Z goes in to the screen)

Representation of 3D Objects

- Purpose
  - For display :- Produce a 2D image
  - For CAD :- Final output is a 3D object
- Solid or surface?
  - Geometrically Based Rendering Systems
    - Wire-frame Representation
    - Polygonal Representation
    - Bicubic Parametric Patches
    - Constructive Solid Geometry (CSG)
  - Volumetric Representation
  - Deformable Modelling
- Static or Dynamic?

Static Model vs. Dynamic Model

- Static Object
  - Single images or objects
  - Trees or walls that don't move
  - Pre-participation in a VE
- Dynamic Object
  - Video like images
  - Object will response to touch and deform accordingly
  - Dynamic object usually rendered by code during participation in VE
- Static Behaviour
  - Fixed or articulated behaviour; behave according to preset parameters
- Dynamic Behaviour
  - Simulated by code during participation in VE

Wire-Frame Representation

- This is the simplest method
- Only the polygon edges are displayed, not the facets
- Low computational complexity
- Earliest method to be implemented in hardware for real time performance
- Suitable for overlaying the wire frame on continuous pixel images in AR applications

Polygonal Representation

- Object surfaces are represented as a set of flat 2-dimensional polygons that combine to form an approximation to the surface. Curved surfaces are decomposed into polygons (tessellated), which are small enough that the individual facets are not visible.
- Advantage is simplicity
- Dominant all surface representation in VR
- Each polygon is uniquely defined by its list of vertex coordinates.
- Geometric manipulations only applies to the vertex hence computational efficient
- The disadvantage is that only the surface at a particular threshold value is matintained.

Polygonal Representation

- 6 polygon are required to model a rectangular box
- Good for solid object and easy to manipulate within a computer
- Polygon are useless for modelling clouds and fog
- We can represent to an accuracy that we choose, an object of any shape
- Final projected size to screen is wastage because the screen area only make up of a few pixels.
- Good for CAD but may not as good for games; need compression

Polygonal Representation

- A hexagon can be easily reduced into a number of triangles.

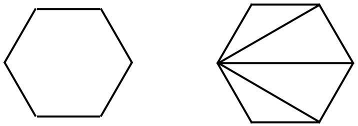


Figure: A single hexagon and a triangulated hexagon

Bicubic Parametric Patches

- The surface is approximated by a 2D piece-wise bicubic parametric patches
- Visually superior to polygons
- Provide more accurate representation of the model.
- An edge list is maintained
- The complexity of any transformation is increased
- http://www.digimation.com/

Constructive Solid Geometry

- The graphic object is expressed as a set of primitive bodies
  - Spheres
  - Cylinders
  - Cubes
  - Cones
  - Tori
- Manipulation using Boolean relationship
  - Union
  - Intersection
  - Substraction, etc.

Volumetric Representation

- The object is represented as a 3D array of “voxels”
- Voxels are volume elements, the 3D version of the 2D pixels
- This method represents the entire volume – solid model
- Maintains the true representation of the desired details
- The techniques used to visualize volumes is known as “volume rendering”

Deformable Modelling

- Virtual objects which bend or deform appropriately when touched or with feedback to heptic devices.
- Geometry-based Deformable Models
- The object or surrounding space is deformed based purely on geometric manipulations
  - Vertex-based or Spline-based
  - The vertices or groups of vertex of the object are manipulated to display the visual deformations. The reaction force can be modelled using Hooks Law, where depth of penetration can be computed based on the current and home positions of the vertex which is closest to the contact point.

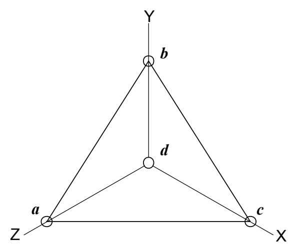
Deformable Modelling

- Physics-based Deformation Models
  - Particle based systems consists of a set of point masses, connected to each other through a network of spring and dampers, moving under the influence of internal and external force.
    - Simulation of soft tissue and cloth behaviour
  - Finite element based systems divides the volume occupied by a 3D object into finite elements, properties of each element are formulated and the elements are assembleeded together to study deformatoon states for a given load. Very computation intensive and simplification is necessary.

Internal Storage of 3D Objects

- Example
- Colour views can be created at a fast speed
- Short delay is important in VR system because the illusion of immersion and presence can get lost quickly.
- Triangle are more preferred to polygon for rigidity
  - 3 sides – the surface is always flat
  - Polygon – minimum 4 sides and the surface can be twisted
  - A polygon can be easily divided into a number of triangles
  - Wildcat 5110 can process 15 million triangle per second (Z-buffered, 25 pixel, Gourad shaded)
- In 3D Cartesian coordinates, each vertex needs 3 measurements to locate it relative to a reference point
- More complicated the object more vertex and it create serious problem for computer. Hence, need to keep the VE as simple as possible.

Internal Storage of 3D Objects



Triangle Table

name	Vertices		
abc	*a*	*b*	*c*
cbd	*c*	*b*	*d*
adb	*a*	*d*	*b*

Vertex Table

name	*x*	*y*	*z*
a	0	0	1
b	0	1	0
c	1	0	0
d	0	0	0

Figure: Internal storage of 3D object

3D Computer Graphics

- Shading algorithms
  - Gouraud 1971
  - Phong 1973
- Texture mapping - James Blinn, 1976
- Anti-aliasing
- Shadows
- Hidden-surface removal
- Environmental mapping
- Modelling strategies

Nongeometric Rendering Systems

- Volume Rendering
  - For semitransparent objects
  - Ray tracing techniques
    - Defining rays of lights as a light source. The ray behave according to the laws of physics as related to light and optics. The light rays are altered as they reflect off of surfaces of the defined virtual objects, taken into account the nature of the material the surface.
    - Computations are very intensive and complex and is not yet generally available in real-time VR operation.
- Particle rendering
  - To show complex flow in a visual scene, many small particles are rendered over time, producing the visual features that reveal the process of a larger phenomenon. For example, combustion process, fire, smoke, water flow and animal group behaviour like a school of fish, a flock of birds or a crowd of people.

Object Shading

- The shading of object surfaces enhances their visual fidelity and simplifies the underlying physics of computing and displaying realistic looking models.
- 2 considerations
  - Accuracy and cost effectiveness
- Shading Algorithms
  - Flat Shading
  - Gouraud Shading
  - Phong Shading

Illumination Models

- An equation that is evaluated at each surface point
- The equation can accept a number of input parameters:-
  - Viewer position
  - Position of surface point
  - Surface characteristics
  - Light-source position
  - Light-source characteristics
- Modern lighting models
  - Ambient
  - Diffuse
  - Specular

Ambient Light

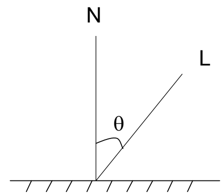
- Approximation to global illumination
  - Each object is illuminated to a certain extent by “stray” light
  - Constant across a whole object
- A constant, non directional light source, often simply used to make sure everything is lit.
- Ambient light thus set for the entire scene (I_a)
- Each object reflects only a proportion of incident light (k_a)

$$I_r = k_a I_a$$

For RGB

$$I_{r,\,red} = k_{a,red} I_{a,red}$$
$$I_{r,green} = k_{a,green} I_{a,green}$$
$$I_{r,blue} = k_{a,blue} I_{a,blue}$$

Lambert’s Law



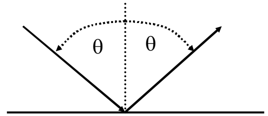
- Reflected intensity is proportional to cos θ
- L is the direction to the light
- N is the surface normal

Diffuse Light

- Reflected from Dull surface such as cloth, etc.
- The surface brightness is proportional to the cosine of the angle between the surface normal and the light direction.
- The normalised intensity of the light incident on the surface due to a ray from a light source
- The light reflected due to Lambert’s law
- The proportion of light reflected rather than absorbed (k_d)
- With both ambient and diffuse components

Specular Light

- Shiny surface exhibit highlights as a result of specular reflection.
- Depends on the viewer’s position as well as the light position
- Phong developed a model that uses the cosine of the angle between the viewer and the light reflected from the source.
- The specular exponent is a measurement of reflection.
- Surface that are good reflectors have a large exponent n.



Other Issues

- Culling
  - Remove invisible pieces of geometry and only sending potentially visible geometry to the graphics subsystem. Simple culling rejects entire objects not in the view. More complex systems take into account occlusion of some objects by others; e.g. a building hiding trees behind it.
- Collision Detection
  - Surround the object by an invisible box called a collision volume which is stored alongside the object's surface geometry in the database

Z-buffer

- Developed by E. Catmull in 1975
- A method for performing visible surface resolution so that only the visible surfaces of an object are rendered.
- The Z-buffer algorithm operates in image (screen) space and associates a z (depth) value with each pixel. The Z-buffer holds the smallest of all z values at the same (x, y).
- At the refresh of screen, new z values for every pixel will be calculated and compared. If the z value is less than the saved value (nearer the viewer), the new intensity and z values replaced the saved values; otherwise, they are discarded.
- The advantage of this method is its simplicity. But it requires additional storage for the z value for each pixel.

Aliasing

- When a straight edge in a binary image is oriented diagonally, it will display "staircasing" or "jaggies".
- Long objects that are thinner than a pixel may break up or disappear completely.
- Dependent upon the orientation of the object related to the screen coordinate axes.
- Object scintillating or Moiré Moiré pattern on texture-mapped surface
  - http://www.exploratorium.edu/snacks/moire_patterns.html
  - http://www.sandlotscience.com/Moire/Moire_frm.htm
  - Moiré pattern can be obtained when you scan newspaper or magazine without descreening
- Undersampling; high resolution information is represented by low resolution means.

Methods to Reduce Aliasing Artifacts

- Supersampling
  - Also known as postfiltering or filtering after sampling
  - The most popular method
  - Sampling is performed at higher resolution (e.g. 8x in each direction) and followed by low pass filtering (e.g. smear, blur, average), then sampling at the screen resolution.
- Prefiltering or filter before sampling
  - Area sampling method by Catmull.
  - The contribution of a pixel to intensity is proportional to the area of interaction between the square pixel and that portion.
  - Sampling after convolving the image with a square box filter
  - Compute the subpixel intensity or use bit masking
  - May be implemented on graphics card and known as antialiased, area-averaged, accumulator buffer (A-buffer).

Methods to Reduce Aliasing Artifacts

- Stochastic Filtering
  - The sample are not equally spaced, trading aliasing for noise.
  - Performing sampling at a sampling grid whose coordinates are perturbed, followed by applying a reconstruction filter to estimate values at the unperturbed sampling grid points.
- Antialiasing for Texture Mapping
  - MIP-mapping
    - mip = multum in parvo = many things in a small place
    - Multiresolution method; scales and filter a texture map into multiple resolution before applying it to the wire frame model.
    - At close distance, use the original, at far away use reduced resolution or smaller maps.

Human Visual System

- 2 types of primary light receptors in the retina
- Rods – function efficiently in dim light
  - Rods contain a purple dye called rhodopsin that is destroyed by light.
  - Takes time to build up rhodopsin or adapted to dark environment
- Cones – function efficiently in bright light
  - Ceased to function below approx. 0.1 cd/m^2
  - Continue to function above 10 cd/m^2
  - Contains iodopsin (3 forms with peak sensitivity in the red / green / blue of the electromagnetic spectrum)
  - Cones can detect colour but not rods
- Population of cones is larger than rods.

Example : Why we see 3D?

- Cues to depth
- Relative image size
- Parallax
- Ocular convergence
- Stereoscopic fusion
- Experience
  - http://www.houseof3d.com/pete/applets/wireframe/stereo/viewmaster.html
  - http://www.jfairstein.com/3d.html

Human Visual System

- Monocular depth cues
  - Motion parallax cues
    - Arise when there is relative motion between the observer and some environment.
    - When looking at a cluster of objects with one eye, any slight head movement expose depth features that remain hidden when stationary.
    - Application – flight simulators
    - The pilot experienced 3D depth when the simulated aircraft is in motion.
    - These display systems can totally encompass the visual field, peripheral vision becomes simulated and complements the pseudo-3D deffect to create an impressive sensation of being totally immersed in a VE

Human Visual System

- Perspective depth cues
  - When we are visually familiar with something, such as the shape of a car, it is possible to estimate accurately their depth. Through the size of the image projected upon the eye's retina.
  - The brain can be easily fooled by staging contrived illusions that incorporate false perspective cues.
- Binocular depth cues
  - Eye convergence / ocular convergence
    - We estimate the depth of the object by changing the lens of our eyes if the object is moving towards us.
    - The convergence angle is zero when our eyes are relaxed and gazing fixation some point on the horizon
    - Objects placed too close to the eyes appear double and cause eye strain

Human Visual System

- Stereoscopic fusion
  - The images on the two retinas are slightly different and the visual cortex of the brain compares the two images and interprets the difference between them in terms of depth. Physiologists call this facility stereopsis.
- Stereoscopic perception varies greatly between individuals, probably only about 5% have the senses developed to the high degree, 20% have only feeble stereoscopic perception and 5% have monocular vision.

Stereopsis

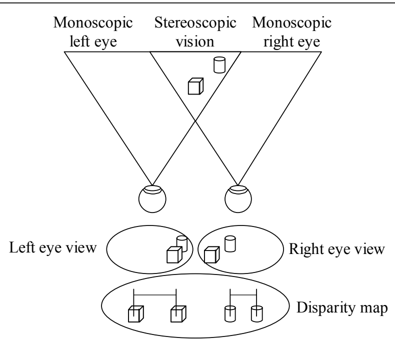


Figure: Diagrammatic description of stereopsis

Virtual Humans

- A VE is not complete without people meeting people or virtual people
- Human modelling
- Construct from reality
- Body deformation is still unsolved
- 4 methods of dealing with facial expressions in Networked VEs:-
  - Video texturing of the face
  - Model-based coding of facial expressions
  - Lip movement synthesis from speech
  - Predefined expressions or animations

3D Presentation Displays

- Mechdyne's rear-projected curved screen is scalable up to a full 360 degrees horizontal field of view (FOV) with a 12-foot (or greater) radius. The solution is available with mirrors to reduce the overall system footprint, and the projection technology utilizes special electronics to accommodate the geometrical corrections required for this type of configuration.
- Sizes vary greatly depending on desired field of view (FOV), radius range 10-50 feet, a standard 12 foot system provides 150 degrees of horizontal and 40 degrees of vertical FOV with a 25' by 20' footprint
- http://www.mechdyne.com/pcsvr.shtml

3D Presentation Displays

- Cathode Ray Tubes (CRT) /Liquid Crystal Displays (LCD)
- Head-mounted Displays (HMD)
  - Helmet-like devices include 3D viewing glasses, stereo audio and tracker devices
    - http://www.virtualresearch.com/products.html
- Shutter Glasses
  - One eye sees the image, the other eye sees black
  - Must synchronize with the display card
  - Wireless models use infrared to connect to the PC
- Projector-based Displays
  - Immersive, wide-angle or peripheral-vision hemispheric
  - http://www.mechdyne.com/pcsvr.shtml
  - http://www.elumens.com

Autostereoscopic Display

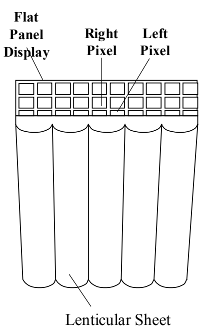


Figure: lenticular display

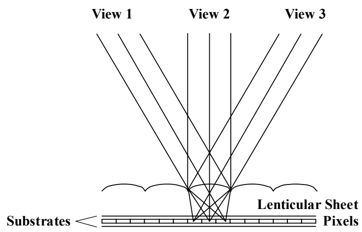


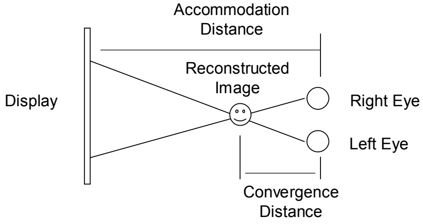
Figure: Three view lenticular display

Example on 3D Display

- One solution is to create the depth in the display
- A LCD monitor product from Deep Video Imaging (a New Zealand Company) – multi-layer display (MLD)
- The MLD monitor is actually two LCD displays stacked on top of each other. The LCD on top displays foreground images transparently, user can see through to a second LCD panel beneath it, which is opaque and display the background images. The two monitors are driven independently of each other via separate video cards; either a dual-head video card or a second video card can be used.
- Example application:- stacking 3D applications for kiosk, multidimensional information display, etc.
  - http://www.deepvideo.com/index3.html
  - http://www.stereographics.com/products/synthagram/synthagram.htm

Problems in 3D Display

- Mismatch between the focus (accommodation distance) and fixation (convergence distance) causes fatigue while watching a stereoscopic TV.
- A major problem in HMD and lenticular lens systems



Input / Output Devices for VR

- Everywhere Display from IBM
  - http://www.research.ibm.com/ed/ed_technology.htm
- Visual display at IBM Research
  - http://www.research.ibm.com/compsci/graphics/index.html
- i-O Display Systems
  - http://www.i-glassesstore.com/index.html
- Head Mounted Displays and other VR display devices comparison
  - http://www.stereo3d.com
- TrackIR
  - http://games.naturalpoint.com/
- VR haptic feedback Interface and application at Sensable Technologies
  - http://www.sensable.com

Display Hardware

- To create immersion, display must fill a wide field of view (FOV)
- Technologies
  - CRT, LCD, Plasma and Projector
- Devices
  - Display close to eyes (HMD)
  - Large display (ImmersaDesk)
    - http://www.evl.uic.edu/pap/CAVE/idesk/
  - Curved display (dome)
    - http://www.virtual-reality.com
  - Multiple displays (CAVE)

Cave Automatic Virtual Environment (CAVE)

- A spatially immersive display (SID) or semi-surrounding projected stereo display where images are projected on the walls, floors and ceiling of a room / booth that surrounds a viewer.
- The advantage of CAVE is it can accommodate multiple people at a time.
- The principal viewer’s head position should be tracked so as to determine the view direction and content.
- Other viewer will suffer if they are not looking into the same direction as the principal viewer.

Rendering Software

- Eye tracking and locate the viewer
- Render stereoscopic view of the VE
- Must work in real time (10Hz – 30 Hz)
- Need low polygon count
- Need simple illumination / shading model (Gouraud)
- Need optimisation (e.g. view culling, LOD)
- Hardware accelerated texture

Interaction Software

- Basic interaction
- Navigation
- Selection / manipulation
- System control

- Interactive Medical Applications
- The visible human project
  - http://www.nlm.nih.gov/research/visible/applications.html
- 3D Virtual Colonoscopy – Animations
  - http://www.cs.sunysb.edu/~vislab/animations/colonoscopy/

VRML

- Virtual Reality Modeling Language
- A script language
- Build 3D virtual worlds on the Web
- Not really a VE because viewer does not immerse in it.
- Animation is interesting
  - http://www.web3d.org/vrml/vrml.htm
- The Humanoid Animation Working Group
  - http://www.h-anim.org/

Aural Display System

- Localization
  - The psychoacoustic phenomenon in which a listener can determine the direction and distance from which a sound emanates.
  - Three-dimensional sound fields need multiple channels.
- Sound stage is the source from which a sound appears to emanate relative to the listener.
  - World referenced sound stage
    - Implemented by loud speakers
  - Head referenced sound stage
    - Implemented by handphones
    - Masks real-world noise

Haptic Displays

- It exists in a form of both input and output.
- Our sense of touch and proprioception is quite powerful.
- Creating a statisfactory display device is difficult.
- Combined sensation of kinesthesia and taction
  - Kinesthesia is the perception of movement or strain from within the muscles, tendons, and joints of the body.
  - Proprioception means stimulation from within the body.
    - An individual's ability to sense their own body posture, even when no forces are acting upon it.

Properties of Haptic Displays

- Kinesthetic cues
  - The combination of nerve inputs sensing the angles of the joint, muscle length, tension and resistance to muscle effort (force).
  - There are 75 joints in the entire body (44 joints in the hands alone) which are capable of receiving kinesthetic cues, making it very difficult for a single display to engage each possible point of force on the user.
- Tactile cues
  - Cues use sensory receptors at the skin to gather input about the world.
  - Mechanoreceptors; thermoreceptors, electroreceptors, nociceptors.
- Grounding
  - The anchor that needs to supply the base against which pressure can be applied; self-grounding or world-grounding.
- Number of display channels
- Degrees of freedom - 6 DOF in unconstrained, free movement.

Properties of Haptic Displays (cont.)

- The form of haptic display
  - a prop
    - A physical object used as an interface to a virtual world; a prop may be embodied by a virtual object and might have physical controllers mounted on it.
- Fidelity
  - For example, the maximum force that can be exerted by a human finger is about 40N, performing precise manipulations rarely exceeds 10N.
- Spatial resolution
  - The brain's ability to discriminate closely tactile stimuli varies by region of the body; on the back, this distance is 70mm, on the forearm, this distance falls to 30mm and on the fingertip, it drops to 2mm.
- Temporal resolution
  - The idea is to eliminate the feel of vibration. The frame rate required for correct force feeling is around 1000Hz. Above 1000 Hz, minimal improvement has been observed.
- Other properties like latency tolerance and the size of the device is depending on application and safety considerations.

Interaction and Manipulation in VE

- Direct User Control
  - Participant interacts with virtual objects as they would on real objects; e.g. touching or grabbing
- Physical Control
  - Control via pressing real button or controller or button mounted on top of the prop
- Virtual Control
  - Screen control, software menu, etc
  - May lose haptic feedback
- Agent Control
  - Voice command

Direction Selection Methods in VE

- Pointer-directed Selection
  - Requires hand tracking but easy to implement
- Gaze-directed Selection
  - Need head tracking; not by eyes focus.
- Crosshair-directed Selection
  - Track both hand and head
- Torso-directed Selection
  - Need the tracking of torso
- Device-directed Selection
  - Use a physical control like a joystick
- Coordinate-directed Selection
  - User input coordinates by using input device such as voice command or keyboard
  - Both relative or absolute coordinates can be used
- Landmark-directed Selection
  - Identify the location / object by tag and input through voice or keyboard.
  - E.g. Big tree, flag pole, etc.

Items Selection Methods in VE

- Contact select
  - Visual, aural or haptic devices; good feedback
- 3D cursor select
- Menu select
- Point-to select
- Name-to select
- Select in mini world
  - A subset of the virtual world is represented as a small model on a palette
  - Similar to contact selection but with specific object
- Aperture select
  - Select on the aperture of fingers, just like pick up something virtually
  - Need finger tracking and head tracking
  - Similar to crosshair selection

Reference

- Sherman, William, and Craig, Alan, “Understanding Virtual Reality - Interface, Application and Design”, Morgan Kaufmann, 2003.
- Burdea, Grigore, “Virtual Reality Technology”, 2nd Ed., John Wiley, 2003.
- Vince, John, “Virtual Reality Systems”, Addison-Wesley, 1995.
- Slater, Mel, and Steed, Anthony, and Chrysanthou, Yiorgos, “Computer Graphics and Virtual Environments – from Realism to Real-time”, Addison-Wesley, 2002.
- Hill, F.S., “Computer Graphics – using Open GL”, 2nd Ed., Pearson Education, 2001.
- Special Issue on Virtual & Augmented Reality in Medicine, Proceedings of the IEEE, Vol., 86, No. 3, pp.469-608, March 1988.