Display & Rendering

029515
2010년 봄학기
3/22/2010
박경신

Human Perception System

- Obtain information about environment through senses:
  - Vision: primary form of perception in most VR
  - Audition: second most common in VR
  - Haptic/Touch: perceptible on through direct contact
  - Vestibular/kinesthetic sense
  - Olfaction
  - Gustation

distant (Vision, Audition) **proprioceptive** (Haptic/Touch, Vestibular/kinesthetic sense) **chemical** (Olfaction, Gustation)

- VR systems mimics the senses by output of computer-generated stimuli rather than natural stimuli to one or more of these sense.

Vision

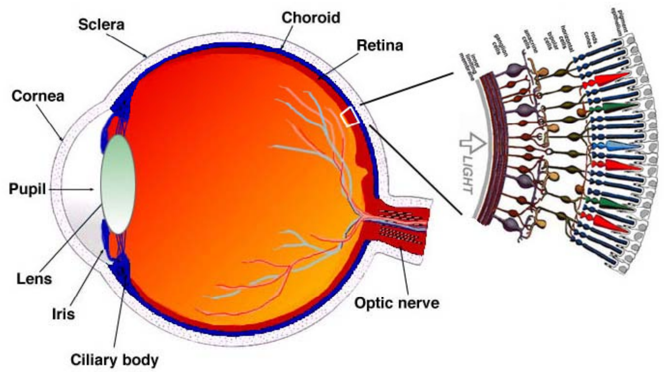
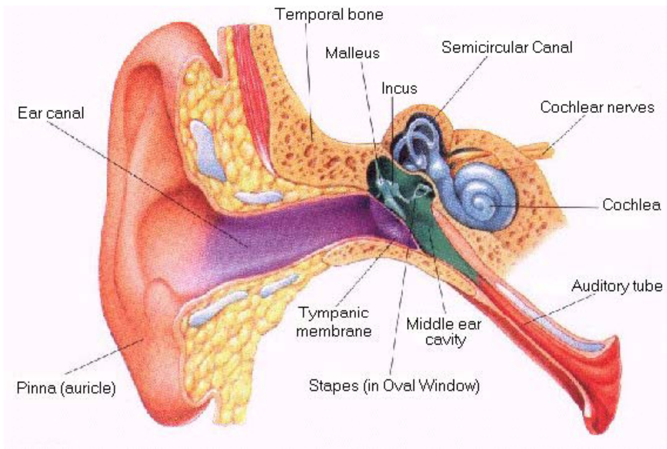
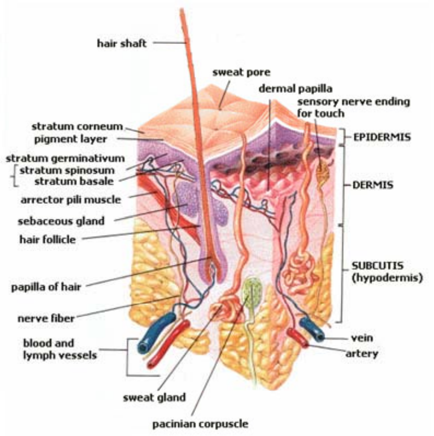


Fig. 1.1. A drawing of a section through the human eye with a schematic enlargement of the retina.

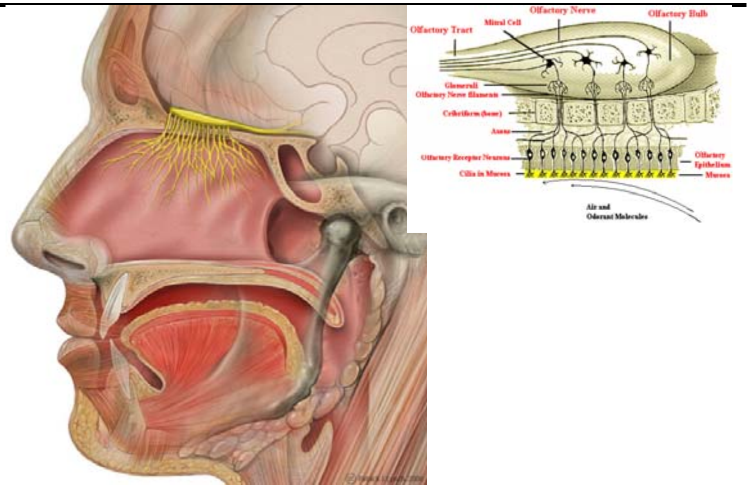
Audition



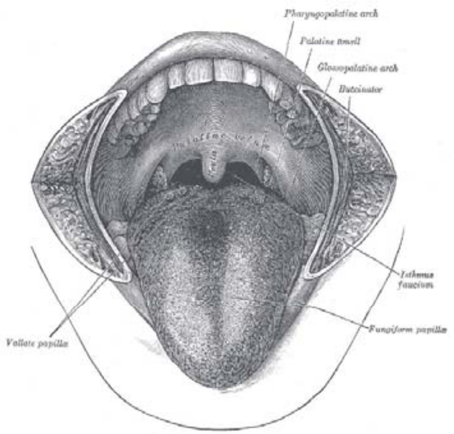
Touch



Olfaction



Gustation



Displays

- Display device presents perceptual information
- Often display is used to mean 'visual display'
- Goal: display devices which accurately represent perceptions in simulated world (i.e., higher levels of immersion)

Visual Displays

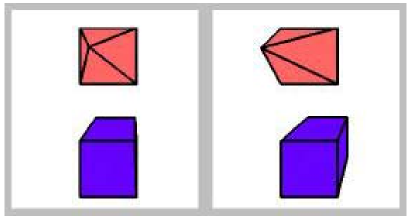
- Fishtank VR stationary display
- Projection VR stationary display
  - Surround-screen displays
  - Tabletop displays
  - Wall displays
- Occlusive head-based display
  - Head-mounted display (HMD)
  - Binocular mono-oriented monitor (BOOM)
- Nonocclusive head-based display
- Hand-coupled display

Visual Display Presentation Qualities

- Color
  - mostly trichromatic color
  - monochromatic color in some displays – e.g. see-through HMD
  - field-sequential color display – overlays the three colors in same location
- Contrast/Brightness
  - dynamic range of the display
  - LCD displays tend to have lower contrast than CRTs
  - See-through HMD requires brighter display
- Spatial resolution
  - number of pixels & pixel density
- Screen geometry
  - Rectangular, L-shaped, hemispherical, hybrids
  - Non-rectangular screen shapes require nonstandard projection algorithm

Visual Display Presentation Qualities

- Number of display channels
  - two for stereoscopic displays
  - sometimes two display channels, but the same image on both
  - many ways to transmit multiple channels: color (anaglyph), polarization, time-multiplexing, spatial-multiplexing
  - can combine techniques to introduce more channels (perhaps for two viewers) – e.g. Fakespace DuoView



Visual Display Presentation Qualities

- Focal distance
  - distance at which images seem to appear
  - typically the screen in stationary displays
  - can be infinite via optics in an head-based display
- Opacity
  - occlude the real world or not
  - CAVE does not occlude the real world
  - most HMDs occlude the real world
  - see-through HMD is generally used for AR applications
- Masking
  - hiding things behind an object
  - a problem when a virtual object comes between the viewer's eyes and a physical object
  - physical objects (e.g. user hand) mask virtual objects in stationary displays

Visual Display Presentation Qualities

- Field of view
  - measure of the angular width of user's vision
  - typical HMDs cover about 100 degree FOV with about 60 degree stereo overlap FOV
- Field of regard
  - amount surrounding space where virtual world is displayed
  - HMDs are typically 100%
  - CAVEs are often much less (except for 6-sided CAVEs)
- Head position information
  - typically position trackers monitor six degree of freedom (DOF) of the participant's head
  - 3-DOF orientation is needed for HMDs
  - 3-DOF location is needed for stationary displays

Visual Display Presentation Qualities

- Graphics latency
  - lag between user movements and the update of the display
  - source of causing nausea or headaches
  - lag is very noticable when rotating head in HMD
  - lag is less noticable when rotating head in CAVE
- Temporal resolution (frame rate)
  - image updates per second (measured as FPS or Hz)
  - motion pictures capture 24 FPS
  - 15 Hz is considered marginally acceptable
  - 10 Hz and below causes brain to notice that it is seeing a series of still images
- Line transfer
  - Front projection, rear projection, laser light directly onto the retina, special optics

Logical Qualities of Visual Displays

- User mobility
  - can effect on mental immersiveness and usefulness of user VR experience
  - e.g. cables that tether the user, tracking systems with limited range, or screens that prevent further physical movement
- Interface with tracker methods
  - type of displays can influence the selection of tracking methods
- Environment requirements
  - conditions of the surrounding space necessary to provide a good VR experience
  - projection-based displays require low light
  - CAVE requires big rooms

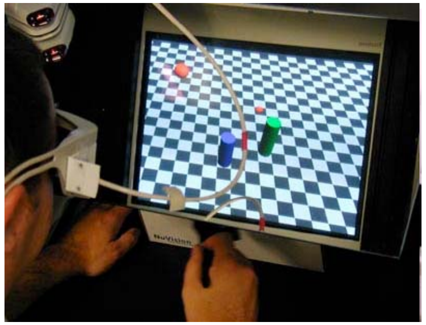
Logical Qualities of Visual Displays

- Associability with other sense displays
  - headphones and HMDs work well together
- Portability
  - large stationary displays are not portable
  - e.g. HMD vs. CAVE

Logical Qualities of Visual Displays

- Throughput
  - HMDs often requires a minute or two to change viewers
  - easier for people to enter and exit a CAVE
- Encumbrance
  - generally more wires associated with HMDs
- Safety
  - eye fatigue and nausea can result from poor optics
  - can't see what you're doing in real world in an occlusive HMDs
- Cost
  - generally head-based displays tend to be lower priced than large-screen projection systems
  - CAVE requires more graphics power

Monitor-based or Fishtank VR

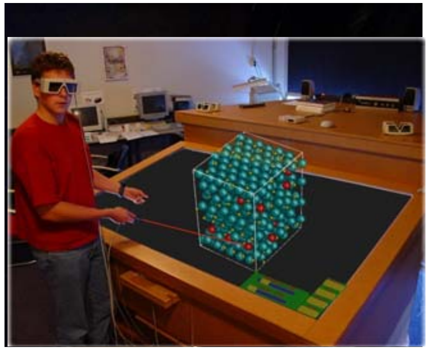


- Use standard computer monitor
- Differ from generic interactive 3D graphics displayed on a monitor because render scenes based on the user's head tracking data
- Fewer components & lower cost
- Standard interface devices available (keyboard, mouse)
- Limited FOV & FOR
- Generally less mentally immersive

Projection-based VR: Surround-screen displays

- Mostly rear-projected to avoid the participants casting shadows on the screen
- Larger more costly displays
- Longer range tracking systems
- Greater FOV & FOR
- Not isolated from the real world
- Multi-viewers friendly
- Not very encumbering
- Less eye strain
- More space required
- More graphics power required
- Occlusion problem
- EVL's CAVE, VRAC's C6, GROTTO, CUBE, RAVE

Projection-based VR: Tabletop displays



- Good for direct manipulation
- Good for god's eye view
- Good for changing orientation
- Less immersion than surrounded VR displays
- Limited mobility than HMD
- Physically-based travel techniques are not suitable
- Responsive Workbench, ImmersaDesk, VersaBench, Barson, VisionMaker

Projection-based VR: Wall displays

- 3D movie-like VR displays
- Larger tiled or curved wall displays are suited for larger audiences – larger pixel, need more projectors
- Less immersion than surrounded VR displays
- Problem of seamless integration
- U. Minnesota's PowerWall, Hemispherical Displays

Head-based Displays

- Small, lightweight screens
- More portable than stationary VR displays
- More encumbering than stationary VR displays
- Distortion at edges
- Tethering to computer
- 100% FOR
- Limited FOV
- No peripheral vision
- Lag in tracking is detrimental
- Eye fatigue

Head-Mounted Displays

- e.g. Sony Glasstron
- HMD Vendors at http://www.faqs.org/faqs/virtual-worlds/visual-faq/section-2.html

See-thru Head-based Displays

- Optical see-through vs. video see-through
- Require 6-DOF tracking
- Registration of tracking with the real world
- Application must live within the restriction of the real world
- Proper rendering of object occlusion is difficult

Handheld VR

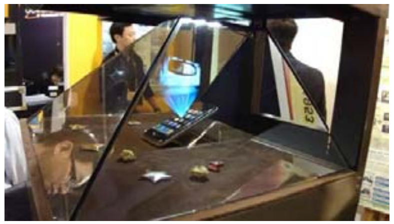
- Limited example of handheld VR – e.g. Chameleon, Virtual binoculars
- Used as a magic lens
- Need to track both the screen and the head
- Registration of tracking with the real world
- Can be used in conjunction with projection-based VR displays

Virtual Retinal Display

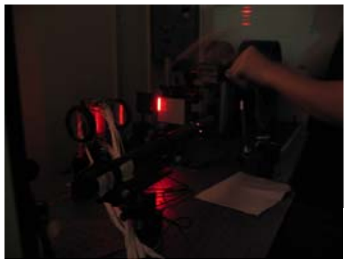
- Also called, light scanning display, by the HIT lab in 1991
- Image scanned directly onto retina
- Viewer can see an image equivalent to a 14-inch monitor viewed from 2 feet away
- High FOV and FOR
- Image loss caused by a lack of eye tracking if users move their eyes while using VRD
- Accommodation and convergence cue conflicts

Autostereo/3D Display

Innovision HoloAD



- 3D Displays
  - Hologram
  - Volumetric
  - Stereoscopic: active, passive, auto stereo, etc



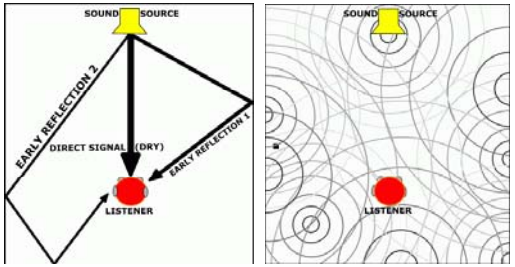
Volumetric Display **Dimen Autostereo Display** **MIT Mark-II Hologram**

Autostereo/3D Display

- Volumetric and Holographic Display
  - Produce true 3D image
  - No accommodation and convergence conflict
  - No motion tracking needed for parallax motion cue
  - Hence, the number of viewers with the correct perspective is unlimited
  - But, current volumetric displays cannot provide many monocular depth cues such as occlusion and shading
  - Volumetric and holographic display can display images only within a small working volume – inappropriate for immersive VE or AR

Auditory Displays

- Second most common VE display
- Perception: pitch, loudness, location
- Technology: speaker-based, headphone-based
- Sound in the real environment:
  - Direct sound: reaches the listener first
  - Early sound: early reflections
  - Reverberant sound: the decay of early sounds



Aural Display Presentation Qualities

- Number of display channels
  - monophonic
  - stereophonic
  - quadraphonic, octaphonic, 5.1
  - multiple speakers rely on ears to naturally localize sounds
- Sound stage
  - source from which a sound appears to emanate
  - **head-referenced** vs. **world-referenced**
    - head-referenced sound stage moves with the head
    - world-referenced sound stage remains fixed with the world
  - sounds require filtering based on head tracking to reproduce a world-referenced sound stage with headphones

Aural Display Presentation Qualities

- Localization (Spatialization)
  - localization is human brain's ability to determine the location from which a sound is emanating
  - spatialization is technology's ability to make a sound appear to come from particular points in space
  - spatialization is easier with headphones due to direct sound control
- Masking
  - loud sounds mask softer sounds
  - physical objects can mask a sound
  - closed headphones are best for VR experience where the participants is only supposed to hear sounds from the virtual world
- Amplification
  - Need to boost the sounds to hearable levels

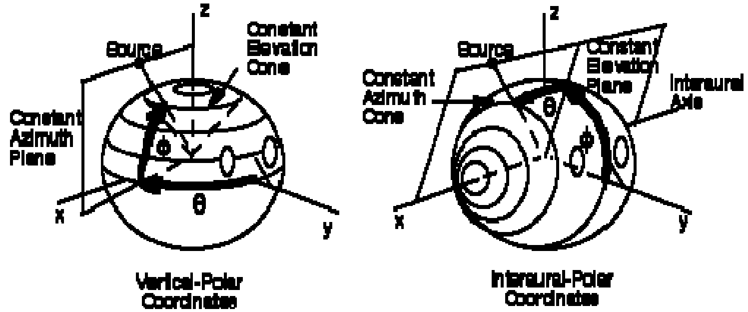
Logical Qualities of Aural Displays

- Noise Pollution
  - speakers require quiet and echo free environment
- User mobility
  - wired headphones limit mobility
- Interface with tracker
  - magnets in the speaker/headphones
- Environment requirements
- Associability with other sensors
  - typically stationary with stationary, head-based with head-based
- Portability & Encumbrance
  - speakers generally more comfortable for longer use
- Throughput
  - speakers work better for larger audiences
- Safety & Cost

3D Sound Localization

- Spatialization
- Works well in plane of ears
- Based on binaural cues:
  - Interaural intensity differences (IID)
  - Interaural time differences (ITD)
- Head-Related Transfer Function (HRTF)

Vertical-Polar Coordinate System



θ: azimuth (angle between the nose and a plane containing the source and the vertical axis z)
ϕ: elevation (angle between the horizontal plane by a line passing through the source and the center of the head)
ρ: range (distance to the source measured along this line)

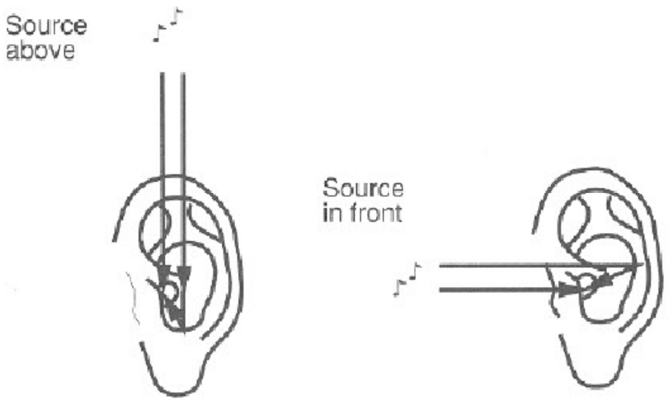
Azimuth Cues

- Interaural time difference (ITD)
  - difference in the arrival time of the sound at the two ears
  - ITD is zero when the azimuth angle is 0 degree, i.e. the source is directly in front of or directly behind the head
  - ITD = $(a/c)(q + \sin\theta)$
    - a: the head radius
    - c: the speed of sound (~343 m/s)
    - θ : source azimuth
- Interaural intensity differences (IID)
  - difference in the intensity of sound reaching the ears
  - the closer ear hears a sound with higher intensity
  - detectable for sounds with high frequencies (>1.5kHz)
  - for low frequency, ITD dominates

Elevation Cues

- Different reflections
  - Due to the asymmetry of the outer ear, especially the pinna
  - the path difference between the direct and pinna-reflected sound changes with the elevation angle
  - sound coming from a source located above the user's head has quite a different reflection path than sound coming from a source in front of the user
- Different amplification (and attenuation)
  - by interference between reflected sounds
  - some frequencies are amplified and others are attenuated
- Pinna provides the primary cue for source elevation
  - user's face and shoulders geometry also influences the way the sound is reflected towards the outer ear

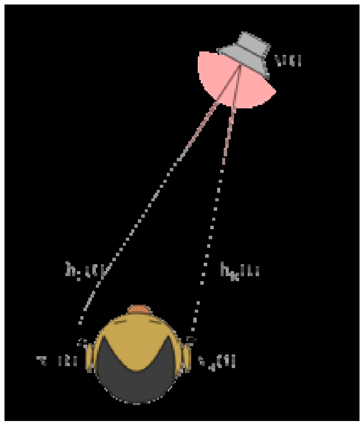
Elevation Cues



Range Cues

- Perceived loudness
  - prior knowledge of a given sound source
  - faint siren (which is a normally high-energy sound source) is perceived as being distant
  - clear whisper (which is a normally faint sound source) is perceived as being close
- Motion parallax
  - change in sound source azimuth when the user is moving
  - large motion parallax indicates a source nearby
- Ratio between direct and reflected sound
  - energy of the direct sound drops off with the square of the source range
  - energy of the reflected sound does not change much with range

Head-Related Transfer Functions

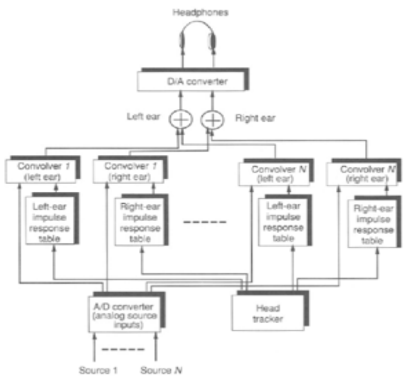


- HRTF captures all of the physical cues to source localization
- Experimental measurement of transfer function
  - sounds from speakers at different locations
  - tiny microphones in the ears
  - analysis of recordings from both ears
  - head-related impulse responses (HRIRs)
  - head-related transfer functions (HRTFs)
- Each individual has his/her HRTF signature, also called ear print

3D Sound Generation

- Two techniques about generating 3D sound
  - 3D sound sampling and synthesis
    - Record sound that the listener would hear in 3D virtual environment by taking samples from real world – But, specific to the environmental settings in which the recordings were made
    - Imitate the binaural recording process by processing a monaural sound source with a pair of left and right ear HRTFs in VE – But, not produce reverberation cues and also many HRTF pairs needed for multiple sound sources
  - Auralization
    - Process of rendering the sound sources to simulate the binaural listening experience through the use of physical and mathematical models
    - Wave-based modeling: solves the wave equation to completely re-create a particular sound field
    - Ray-based modeling: ignores the wavelengths of sound waves and only considers the paths taken by the sound as they travel from source to listener

Convolvotron



- Crystal River Engineering
- HRTF-based spatial audio system
- The system can be customized for a particular individual by measuring and using that person's HRTF
- Echoes and room reverberation can be added by including a room simulation model
- Head motion can be accounted for by combining the absolute location of the source with the outputs of a head tracker to select the appropriate HRTFs

Ambisonics

- Presentation of 3D spatialized sound using multiple stationary speakers
- surrounding sound recording, synthesis and playback system

http://en.wikipedia.org/wiki/Ambisonics

Simple Virtual Environment Audio

- Intensity fall-off & localization
  - 3D spatial sound creates an important audio depth cue
- Headphones also block real-world noises
- Ambient sound in the background
  - Ambient sound effects provide a sense of realism in VE
- Subwoofer in seat or platform
- Present speech instead of text
  - Recorded or synthesized speech can play a role as an annotation tool, and provide help to users when interacting in VE
- Sensory substitution
  - E.g. a sound could substitute for the feel of a button press or physical interaction with a virtual object

Haptic Displays

- Kinesthetic/Force displays
- Tactile displays
- End-effector displays
- Robotically operated shape displays

Haptic Display Presentation Qualities

- Kinesthetic cues
  - nerve inputs that sense angles of joints, muscle length, tension, and resistance to muscle effort (force) within the body
  - helps us determine firmness, approximate shape, and physical force
  - Active kinesthetic vs. passive kinesthetic
- Tactile cues
  - sensory receptors at the skin
  - mechanoreceptor – shape and surface texture
  - thermoreceptor – heat
  - electroreceptor – electric current flow
  - nociceptor – pain
- Grounding
  - force/resistance displays require an anchor
  - self-grounded vs. world-grounded

Haptic Display Presentation Qualities

- Number of display channel
  - how many points of contact with the body
  - 1 channel when Phantom has one point where the user can influence the virtual world
- Degrees of freedom
  - 6-DOF in unconstrained movement
  - 1-DOF display for how far can the thumb be opened/closed
  - 1-DOF display for how far down a tube can you insert a laparoscope camera
  - 2-DOF display for how far down a tube, plus twist
  - 3-DOF display for down, twist, clamping action and resistance
  - 3-DOF display for location of the finger or stylus
  - 6-DOF display for location and orientation

Haptic Display Presentation Qualities

- Form
  - the shape of the physical unit with which the user interacts
  - generic form, such as stick, ball, or plane
  - specific object, such as handgun, or steering wheel
  - amorphous that changes shape to multiple specific representations
- Fidelity
  - how rapidly the system can change to the proper display (force, temperature)
  - can be rated by a maximum stiffness measurement taken in Newtons/meter (Nt/m)
  - a stiffness of 20 Nt/cm as a solid immovable wall
  - 40 Nt/cm is the maximum force that a human finger can exert
  - 10 Nt/cm is the highest force used when doing fine manipulation

Haptic Display Presentation Qualities

- Spatial resolution
  - higher resolution required at the fingertip
  - fingertips can sense difference 2mm apart
  - 30 mm on the forearm & 70 mm on the back
- Temporal resolution
  - how quickly the system can be updated to new display
  - low frame rate on a force display causes the object to be perceived as shakey
  - 1000 Hz is a good minimum
- Latency tolerance
  - low latency display is crucial, especially for force display
- Size
  - larger displays allow broader range of motion

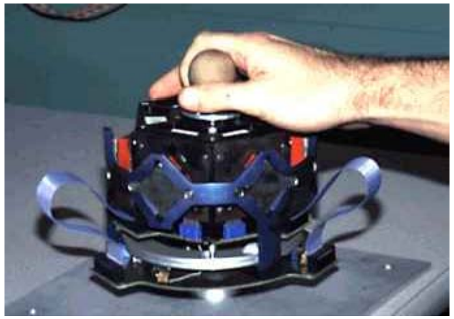
Logical Qualities of Haptic Displays

- User mobility
  - world-grounded displays require the user to be near the device
- Interface with tracker
  - responsive and accurate tracking system is required
- Environment requirements
- Associability with other sense displays
  - occlusive HMDs often are used in conjunction with haptic displayss
- Portability
- Throughput
- Encumbrance
  - self-grounded, exoskeleton-style devices are generally much encumbering
- Safety
  - safety is a significant concern with many haptic displays

Pen-based Haptic Displays

SensAble Tech. Phantom

Hand-based Haptic Displays

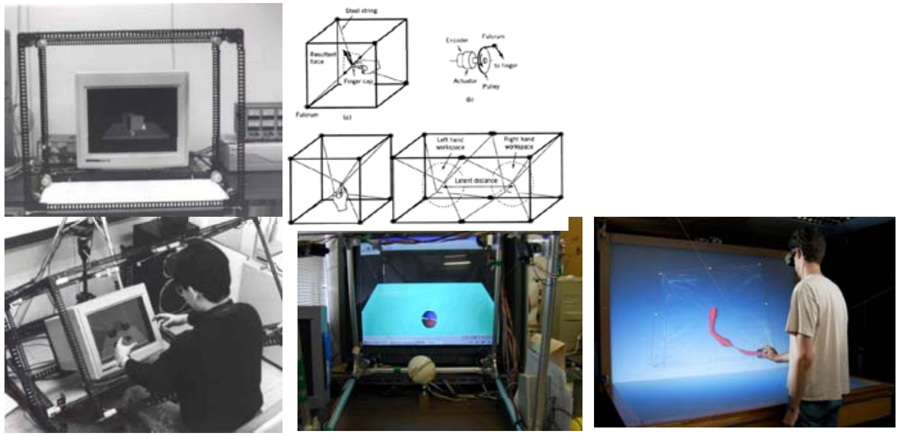


Magnetic Levitation Haptic Interface, Robotics Institute, CMU
Magic Wrist & UBC Wrist, 6-DOF, 20N, 4.5 mm motion range, less than 5 μm

Hand-based Haptic Displays

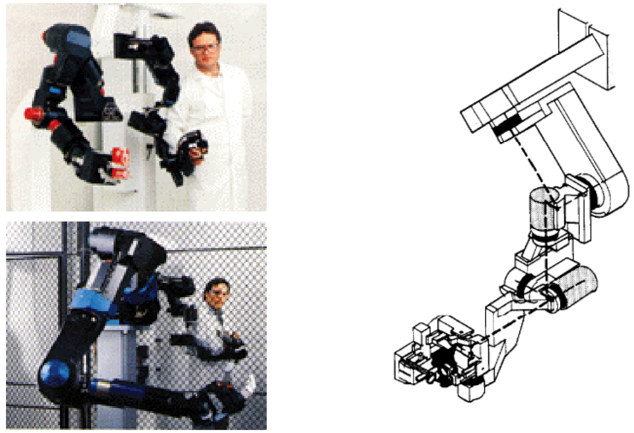
Immersion Corp., Laparoscopic Surgical Workstation

String-based Haptic Displays



Space Interface Device for Artificial Reality (**SPIDAR**) can measure end-point in 3D space & display reflect force.

Arm-based Haptic Displays

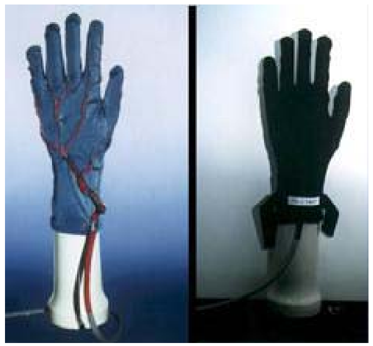


Sarcos' Dexterous Arm, includes a human-sized slave that is commanded by a master system.

Exoskeleton (Body-based Haptic Displays)

The BLEEX project, UC Berkeley, a self-powered exoskeleton, provides a versatile transport platform

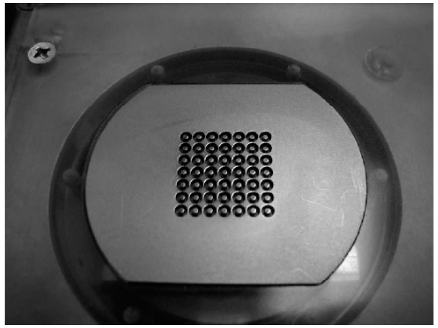
Tactile Displays



Teletact Glove

- Sensed by the skin – the largest single organ of the human body
- Actuators mounted generally on the fingers and hand
- Generally no need for world grounding
- Bladder actuators
- Vibrator actuators
- Pin actuators
- Thermo actuators
- Helps in the fine manipulation of virtual objects
- Less expensive & portable

Electro-tactile Displays



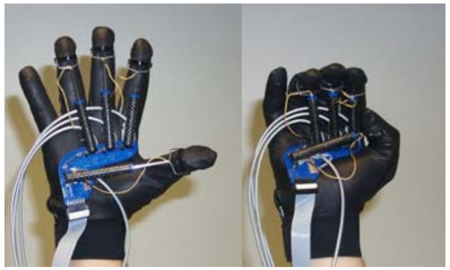
Kaczmark, Electrode array scanned by the fingertips of participants

Tactile Displays using Vibrators

CyberTouch (based on Cyberglove)

Cricket Prob

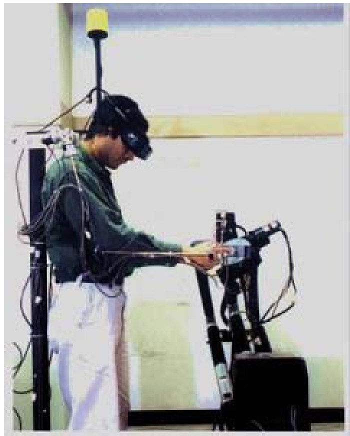
End-effector Displays



Rutgers Dextrous Master

- A mechanical device that provides a force to the participant's extremities
- Generally linked to mechanical tracking sensors
- Generally world grounded (exoskeleton method is body-grounded)
- Often operate with respect to a single point in the virtual world
- Fast and accurate tracking is usually built into display

Robotically Operated Shape Displays



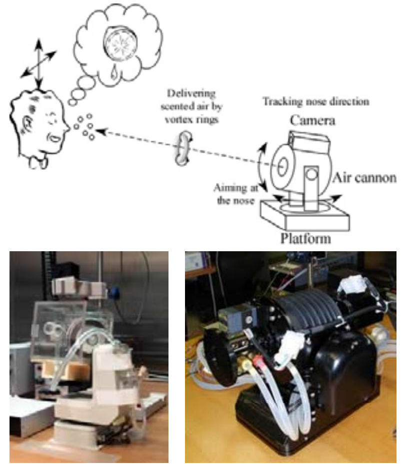
Cybernetic Systems

- Use robots to place a representation of the virtual world where the user is reaching
- May be generic (corners and edges) – e.g. Cybernetic Systems
- May be specific (selection of switches) – e.g. Boeing
- Usually uses a finger surrogate for fast tracking
- Can provide a very realistic haptic display
- World-grounded display
- Works with HMDs

Olfactory Displays

- Very little research is done in olfaction
- Lack of effective displays and difficulty in producing broad range of stimuli
- Olfactory events (odor sources) may be near or far, but directional sensitivity is generally poor
- Temporal sensitivity is poor and response times are slow
  - May need 20-60 seconds between stimuli to resolve different smells
  - To control over stimulus decay rate (without significant air circulation)
- Smell synthesis
- Require chemicals
- Olfactometer
- Smell-O-Vision

Olfactory Displays



- ATR Scent Projector
  - Projection-based olfactory display with nose tracking
  - Unencumbering: Users do not need to wear any devices or glasses
  - Localized: Scent can be perceived only within a limited range of space at a certain time
  - Composed of "air cannon", scene generator (aroma diffuser), 2DOF platform and a CCD camera

Image from http://www.mis.atr.jp/past/sem/scent.html

Olfactory Displays

- Wearable Olfactory Display by U. of Tokyo
  - odor-generating unit (air-pump & odor filters)
  - odor-controlling unit (a notebook PC and a device controller)
  - odor-presenting unit (to present the mixing of odor air)

Gustatory Displays

Food Simulator

- Affected by other senses – strong influence of smell on taste
- Need more than flavor – e.g. texture
- Basic elements of taste – salt, sour, bitter, sweet, umami, smell

Vestibular Displays

- Physically move the user – e.g. motion platform
- Sense of body movements or acceleration
- Vestibular information works together with visual and kinesthetic information
- Virtual body representation
- Can "display" to these senses by stimulating the proper parts of the brain

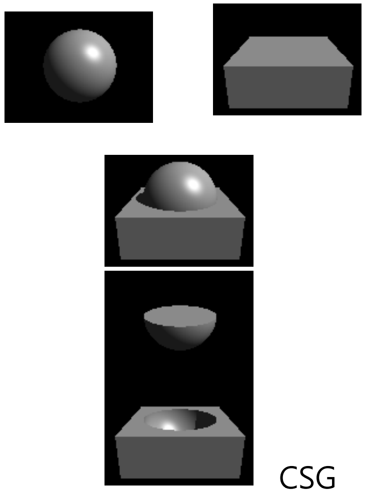
Rendering Systems

- Visual
- Aural
- Haptic

Visual Rendering Systems

- Computer graphics
  - generating visual imagery
- Software rendering
  - graphical rendering routines
- Object presentation schemes
  - Geometrically based (polygons, NURBS, CSG)
  - Non-geometric forms (volumetric rendering, particle systems)

Geometrically-based Representations



CSG

- Polygons
- Constructive Solid Geometry (CSG)
- Non-Uniform Rational B-Splines (NURBS)
- Other representations are often converted to polygons for hardware rendering

Non-geometrically-based Representations

- Volume rendering
- Particle systems

Techniques for rendering complex visual scenes

- Shading
- Reducing polygons:
  - Texture mapping
  - View culling
  - Level of Detail (LOD)
  - Atmospheric effect, e.g., fog
- Multiplexing multiple renderers to one screen:
  - Added cost of additional rendering systems
  - Decrease average image latency for each frame
  - Does not decrease onset latency
  - Reduces the maximum delay between the input and the response

Internal Computer Representation

- Most hardware graphics rendering engines are optimized for polygonal representation
- 3D models created by a model package, Alias, SoftImage, AutoCAD, VRML
- A scene graph is a mathematical graph that allows objects and object properties (colors, materials, textures) to be related to one another in a hierarchical fashion.

Aural Rendering Systems

- Sampling
  - A common way of producing sounds
  - Playback of digitally recorded samples of physical world sounds
  - 8 KHz telephone, 44 KHz CD, 96 KHz DVD quality
- Sound synthesis
  - Spectral method using sound wave's frequency spectrum
  - Physical model using physics of the object generating sound
  - Abstract synthesis is to create sound using some numerical system

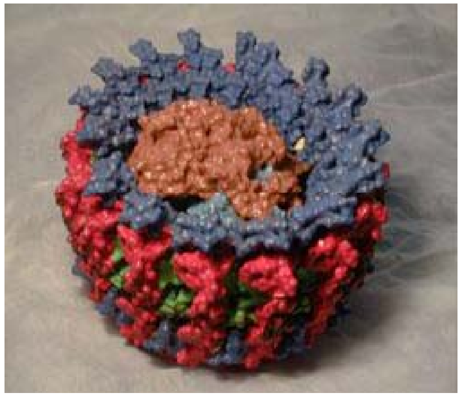
Techniques for Rendering Complex Sounds

- Frequency modulation (FM)
- Algorithmic additive and subtractive techniques
- Granular synthesis
  - Combining the sound of a single drop of water falling to produce the sound of a waterfall
- Sonic effects
  - Convolution – making a sound appear to come from a particular location
  - Reverberation – using reflections of the sound
  - Chorusing – mixing sounds

Haptic Rendering Systems

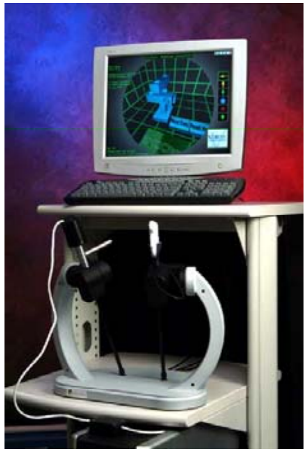
- Thermal rendering – usually used on Peltier thermoelectric coolers
- Pin-based rendering
- Kinesthetic rendering using force display
- Robotically operated shape display
- Physical object rendering -3D hardcopy, aka stereolithography

Stereolithography



http://en.wikipedia.org/wiki/Stereolithography

Techniques for rendering complex haptic scenes with force displays



Laparoscopic surgery interface

- Single point of contact with an object
  - E.g. to a fingertip or tip of a stylus
  - Required 3-DOF force display
- Single point of contact with torque
  - Required 6-DOF force display
- Constraint of movement
  - E.g. laparoscopic
- Two points of contact (pinching)
- Multiple points of contact (grasping)

Haptic Rendering Techniques

- Spring and dashpot
  - controls direction, tension, and damping
- Point and plane & Multiple plane
  - interaction between a probe stylus and a surface by placing a virtual planer surface tangential to the probe's tip
- Point to point
- Multi-springs
  - adds torque to any of the other haptic representations
- Inertial and resistant effects
  - resistive forces to add friction and viscosity
  - inertial force to add momentum
- Vibration
  - a signal indicating when the display should vibrate and at what frequency and amplitude

Reference

- An Interactive Introduction to Splines http://www.ibiblio.org/e-notes/Splines/Intro.htm
- Haptic Community Web Site http://haptic.mech.northwestern.edu/
- 3D audio http://www.dcs.gla.ac.uk/research/gaag/dell/report.htm
- VR audio http://vrlab.epfl.ch/~thalmann/VR/VRcourse_Audio.pdf
- Scent http://www.mis.atr.jp/past/sem/scent.html
- Image from http://webvision.med.utah.edu/sretina.html
- Image from http://www.infj.ulst.ac.uk/~pnic/HumanEar/Andy's%20Stuff/MScProject/workingcode_Local/EarChapter.html
- Image from http://en.wikibooks.org/wiki/Human_Physiology/Print_Version
- Image from http://en.wikibooks.org/wiki/Human_Physiology/Print_Version
- Image from http://en.wikibooks.org/wiki/Human_Physiology/Print_Version
- Image from http://www.siggraph.org/s2003/conference/etech/food.html