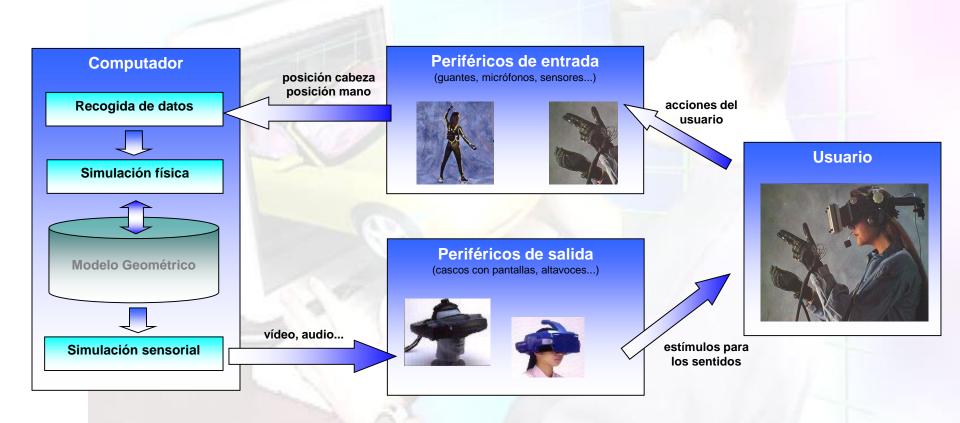
VIRTUAL REALITY

Input and output devices

Course 2024/2025

Architecture of a VR System



Motivation

- Virtual Reality systems use, apart from conventional hardware, a specific hardware which is worth to know.
- In general this hardware is not well known, but:
 - It spends a big part of the cost of a complete VR system.
 - The success or not of the system can depend on the correct choosing of these devices.
- ➤ Importance of the training in VR input/output systems.

Overview

- 1. Output devices
 - a) Visual display characteristics
 - b) Visual display systems
- 2. Input devices
 - a) Concepts and classification
 - b) Obtaining position (trackers)
 - HMDs + tracking
 - c) Data Gloves
 - d) Kinect
- 3. Stand-alone systems
- 4. Devices ad-hoc



Overview

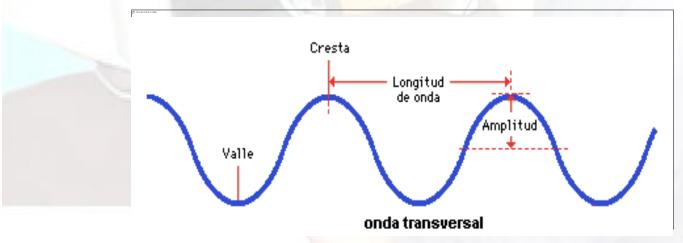
- 1. Output devices
 - a) Visual display characteristics
 - b) Visual display systems
- 2. Input devices
 - a) Concepts and classification
 - b) Obtaining position (trackers)
 - HMDs + tracking
 - c) Data Gloves
 - d) Kinect
- 3. Stand-alone systems
- 4. Devices ad-hoc

Human senses

- Five classical senses
 - Sight (vision) light perception
 - Hearing (audition) sound perception
 - Touch (mechanoreception) pressure perception
 - Taste (gustation) chemical perception
 - Smell (olfaction) chemical perception
- Other senses:
 - Equilibrioception balance and acceleration (vestibular sense)
 - Proprioception (kinesthetic sense) perception of where the various parts of the body are located at any one time.
 - Other internal senses: thermoception, pain, hunger, thirst...

Background: light as a wave

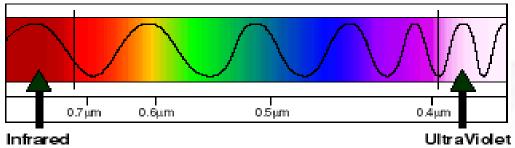
- Visible light is electromagnetic radiation.
- Light propagates through transverse waves.
- Electromagnetic waves magnitudes:
 - Frequency (cycles per time unit)
 - Wavelength: distance between two sequential crests (nm).
 - Visible light: 350 nm (violet) 760 nm (red).



Background: light color

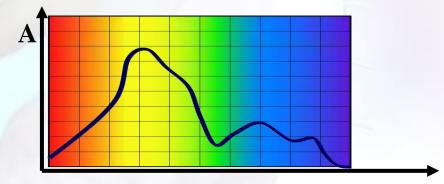
Color of monochromatic light depends on frequency:





The light we see is the result of superimposing multiple waves with different frequencies.

Color of light → spectral diagram

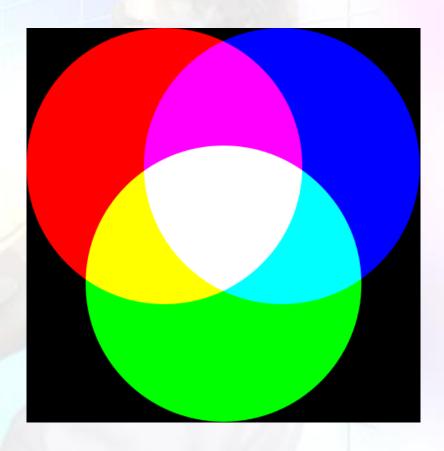


Background: Additive theory (RGB)

Values R, G & B into [0,1]

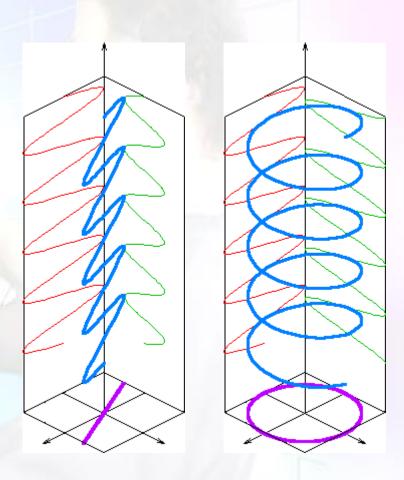
R	G	В	Color
0	0	0	Black (absence of light)
1	1	1	White (max light)
1	0	0	Red
0	1	0	Green
0	0	1	Blue
1	1	0	Yellow (red + green)
1	0	1	Magenta (red + blue)
0	1	1	Cyan (green + blue)

Any color perception can be produced by combining appropriate amounts of red, green and blue light.



Background: polarization

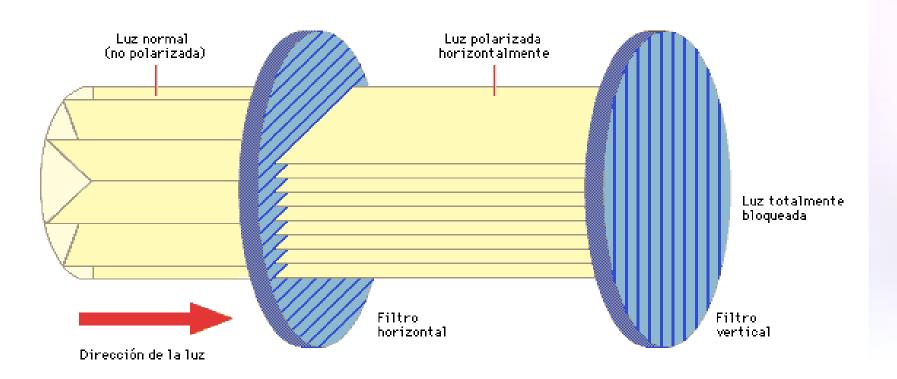
- Polarization: plane where the electric field vector is propagated.
- Non-polarized light: each wave in an arbitrary plane
- Linearly polarized light: all waves in the same plane
- Circularly polarized light: the electric vector rotates either clockwise or counterclockwise



Background: polarization filters

- Linear polarizer: polarization axis
- Circular polarizer: CW/CCW

Transmitted energy



Visual display technologies

For screens:

- CRT Cathode Ray Tube
- LCD Liquid Crystal Display
- PDP Plasma Display Panel
- LED Light Emiting Diode

For projectors:

- CRT Cathode Ray Tube
- LCD Liquid Crystal Display
- DLP Digital Light Processing
- Laser projectors









More information in extra slides

Visual display characteristics

- Field of regard and field of view
- Spatial resolution
- Screen geometry
- Light transfer mechanism
- Refresh rate
- Luminosity
- Ergonomics

Field of regard / field of view

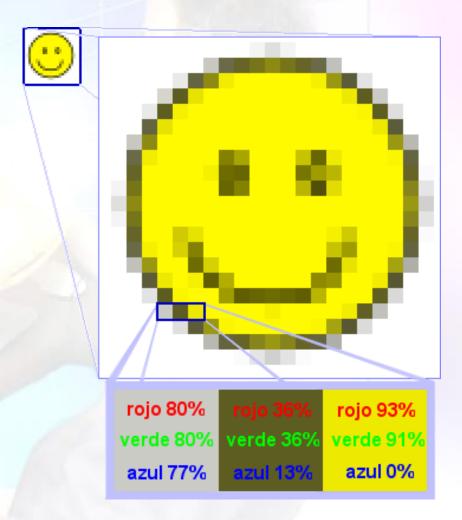
- Field of view (FOV): amount of visual angle that can be seen simultaneously on a display.
 - Display FOV <= Human FOV (200 degrees hor)
 - Example: HMD: 40 to 110 degrees hor
- Field of regard (FOR): amount of space surrounding the user in which visual images can be seen.
 - Example: HMD: 360°
- FOV <= FOR
- Both can be measured in horiz, and vertical directions
- Impact on immersion and stereo perception

Resolution

- Resolution can refer to
 - The number of pixels (eg 1600x1200)
 - The number of pixels per area (eg 100 dpi)
 - The number of pixels per visual angle (angular resolution)
- Human eye → ~ 30 cycles/degree

Resolution

- Quality depends on the number of pixels per area of the screen
- Each pixel keeps information of color (RGB/RGBA)
- Ultra HD (4K): 3840 x 2160
 not enough for a wall of 7 m



Color depth

- Number of different colors
 - 16,7M (24 bits), 4.294M (32 bits)
- 8 bits R +
 8 bits G +
 8 bits B =
- 24 bits per pixel
 → 2²⁴ = 16.7 M
 colors
- CRT, LCD, PDP, LED: unlimited number of levels (theoretically)
- DLP: limited by mirror frequency

Screen geometry

- Number of screens (1-6)
- Shape of the screens (planar, cylindrical, hemispherical)
- Location (L-shaped, tiled displays...)
- Screen dimensions

Light transfer

- How the light actually gets transferred onto the display surface (if any):
 - Front projection
 - Rear projection
 - Special optics
 - No display surface (directly onto the retina)
- Dictates which 3D UI techniques can be used.

Refresh rate

- Refresh rate (Hz) device dependent
 - Number of times the image is refreshed
 - Different concept on CRT/LCD/DLP
 - Low refresh rate → flickering
- Frame rate (fps) application dependent
 - Number of times a complete frame is drawn in the FB
 - Low frame rate → non-smooth animations

Visual displays: systems

- Virtual reality
 - Monitors
 - Head-mounted/coupled displays (HMDs)
 - Projection-based displays
 - Virtual retinal displays
 - Autostereoscopic displays

HMD and HCD

- Do not allow the perception of the real environment.
- Use tracking sensors
- Types:
 - HMD (Head-Mounted Displays)
 - Not attached (except cables)
 - HCD (Head-Coupled Displays)
 - Attached to a mechanical structure
 - Accurate tracking







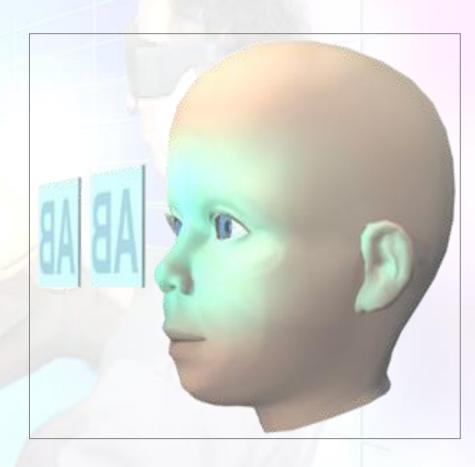




HMD and HCD







Projection-based displays

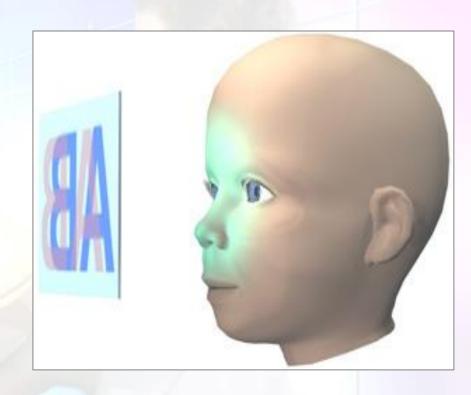
Differ on the number, shape and position of the screens:

- Powerwall
- Workbench
- Holobench
- Reality Rooms
- Domes & spherical displays
- CAVE

Projection-based displays

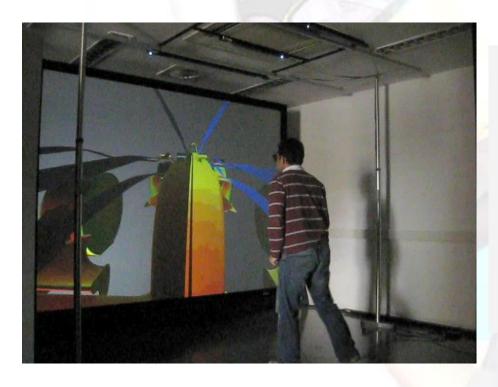






Powerwall, CAD-wall

- The Powerwall was one of the first multi-channel walls.
- Uses projector arrays (e.g. 2x2, 3x2...)





Workbench, ImmersaDesk

It can be used as a virtual tabletop.









Holobench, Barco Consul

L-shaped configuration to let the virtual object protrude further from the screen





Curved screens, reality rooms

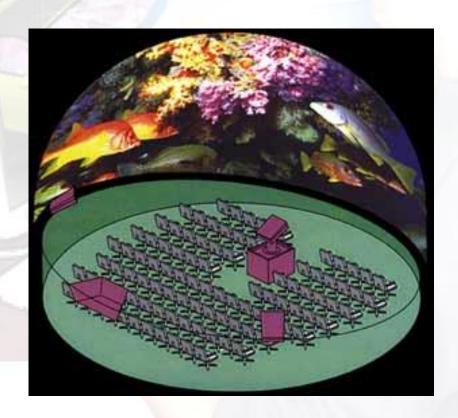
- Mono configuration: used for various simulators (car, cockpit, ship's bridge,...):
- Stereo: less common due to multiple handicaps





Domes & spherical displays

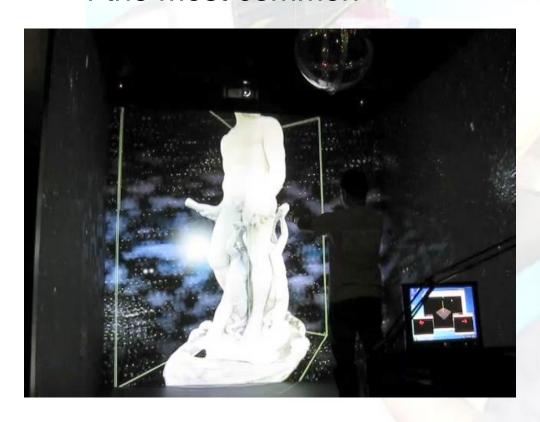
- This configuration is most often found in aircraft simulation and planetariums.
- Challenging due to the multi-sided blending required.



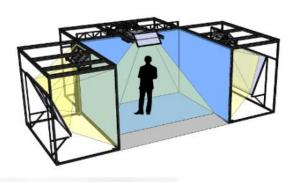
CAVE

From 3 to 6 walls

4 the most common







Advantages of Projection-based

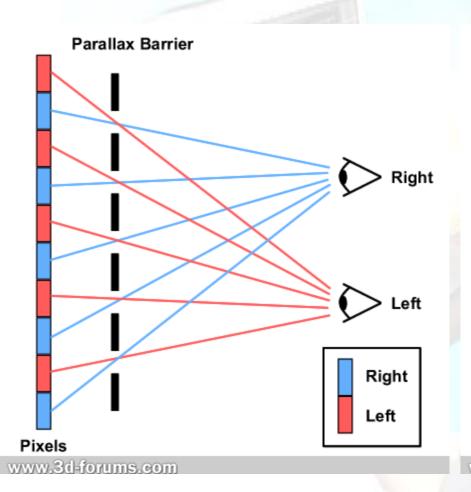
- Ergonomics: the user does not have to wear the display on his/her head
- Do not isolate from real world; do not hinder interaction with other users
- Users can see their own body (HMD-disorientation problem)
- HMD head-rotation instability (flight simulator sickness)
- Collaboration with other users

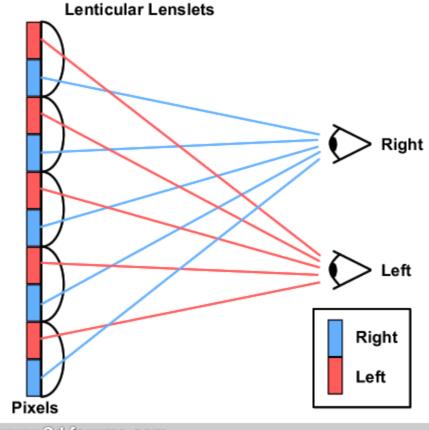
Autostereoscopic displays

- Require no special glasses to achieve stereo.
- Work through parallax separation: the image is split into odd and even columns, and optics ensure that each eye only sees one group.

Autostereoscopic displays

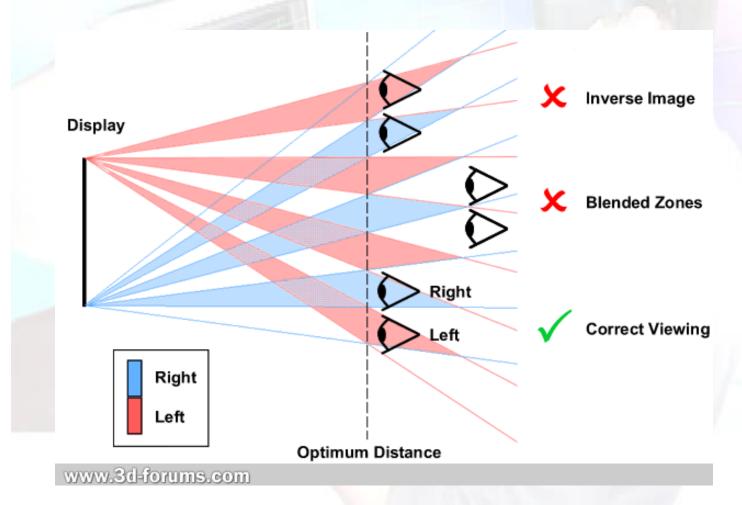
Parallax barrier vs Lenticular array





Autostereoscopic displays

Optimum distance and location





Overview

- 1. Output devices
 - a) Visual display characteristics
 - b) Visual display systems
- 2. Input devices
 - a) Concepts and classification
 - b) Obtaining position (trackers)
 - HMDs + tracking
 - c) Data Gloves
 - d) Kinect
- 3. Stand-alone systems
- 4. Devices ad-hoc

Input devices

- Input devices (**sensors**): capture the *participant* actions (for example the head movements) and send this information to the computer which is in charge of the interactive simulation.
- They are considered as virtual reality devices when:
 - They use the paradigm of the implicit interaction, or
 - They give 3D input to the system.

Classification

- Implicit interaction:
 - Trackers
 - Data gloves
 - Voice recognition
 - Kinect
- 3D Input:
 - Mouses 3-6DOF.
 - Space balls, etc.

3D input devices

3-6 DOF:

- With base surface:
 - Space ball: displacement and angular moment
- Without base surface:
 - 3D balls, Stylus, 3D joysticks...









Trackers

Objective

 The tracking input systems are sensors which are in charge of capture the position and/or orientation of a real object.

Possibilities

- Only position (3 DOF)
- Only orientation (2-3 DOF)
- Position and orientation (6 DOF)



Applications

- Head-tracking, eye-tracking
 - Images eye-referenced
 - Navigation based on head movements



Applications

- Hand-tracking, device-tracking
 - Direct manipulation of objects (multiple metaphors).
 - Navigation hand based (together with gloves).
 - Signs language. Communication with other users.





Applications

- Face-tracking, Body-tracking
 - Interaction based on avatar (games)
 - Capture off-line of complex sequences (sports)





About position

The position can be done by two ways:

- Absolute position (x, y, z): the position is given in known units with respect to a fixed coordinated system.
- Relative position (dx, dy, dz): the position is given by a displacement with respect to the initial position or to the one before.
 - Require calibration (should be in a known position at the initial).
 - The error is accumulated.
 - Non exact.

About orientation



- Two reference systems are defined:
 - The global one (fixed, referenced for example to the emitter)
 - The local one (dynamic, referenced to the object)
- We are interested on knowing the rotation (GT matrix) which transforms the global RS in the local RS.
- The rotation can be represented as follows:
 - Direction cosines
 - Euler angles
 - Angle axis representation
 - Spherical coordinates, polar coord
 - Quaternion

More information about rotation techniques in extra slides

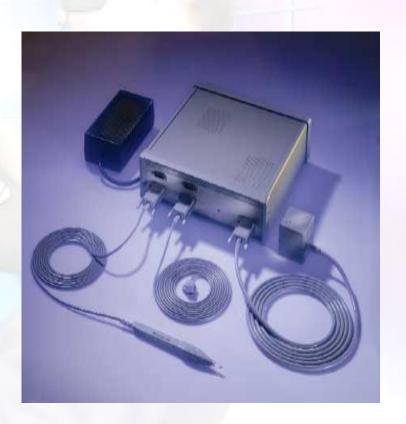
Technologies

Technology of trackers:

- Magnetic
- Acoustic
- Mechanical
- Innertial
- Optical
 - · Outside-In
 - Inside-Out

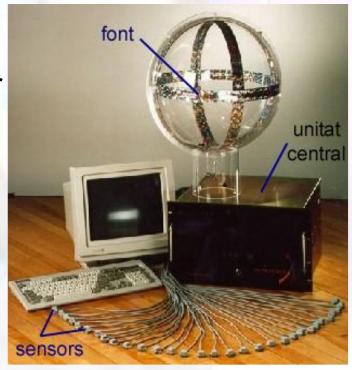
How they work:

- Are based on the attenuation of the orthogonal magnetic fields.
- Each system includes:
 - Emitter: generates the magnetic field.
 - Receptors: detect the magnetic field.
 - Control unit.



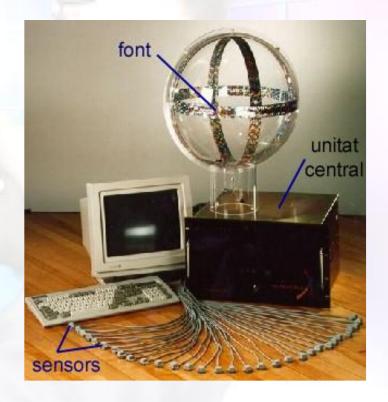
Emitter:

- Is fixed. Defines the global RS.
- It has three orthogonal bobines through which the electric power is passing in a rotative way.
- Each bobine produces a magnetic field perpendicular to the others (sequential in time).



Sensor:

- They are fixed to the object you want to follow.
- They have 3 bobines, but in this case the mesurement is the electrical power induced by the magnetic field which envolves the sensor.
- They are usually tied to the central unit by a cable.



Control unit:

- Gives power to the emitter
- Read data from the receptors
- Solve an equacions system:
 - Variables: 6 (3 position + 3 orientation)
 - Equations: 9 (3 lectures/field * 3 fields)
 - Is undetermined (hemisphere).
- Send the data (via RS-232, for ex.)



Advantages

- Do not require "line-of-sight"
- Very precise.
- Multiple receptors; small.

Inconvenients:

- Area of action limited (2 and 10 meters)
- Sensitive to ferromagnetic surfaces
- Sensitive to other magnetic fields
- Cables

How they work:

- Based on mesuring the propagation time (timeof-flight) of acoustic signals (ultrasons) generated by loudspeakers and registered by mycrophons.
- Different sort of:
 - Fixed emitters, mobile receptors
 - Mobile emitters, fixed receptors

- Several combinations:
 - A mobile emitter and three fixed receptors:
 - Enough to calculate position (3 DOF).
 - Intersection of three spheres.
 - Three emitters and three receptors:
 - Enough to calculate position and orientation (6 DOF)
 - The orientation is trivial from the 3 points (for ex. origin, point in X axe and point in Y axe).



- Advantages
 - Very reduced cost
- Inconvenients
 - Line-of-sight
 - The signals are distorted very easily.
 - Sensitives to ambient noise, eco...
 - Latency proporcional to the distance (ex. 3 meters, 10ms (minimum).
 - Not very precise.

Mecanical trackers

How they work

- Articulated structure with potenciometres attached to the joins.
- Each potenciometre mesure the angle of rotation.
- The base point is fixed.
- The other side point is free
- The position and orientation of the free part: concatenation of translations and rotacions



Mecanical trackers

Advantages

- Fast (almost no latency)
- Very precise
- Electronic very simple.



Inconvenients

- Require to fix the object to the free part of the structure.
- Limits the freedom of movements.
- Cost depends on the mechanical structure.

Innertial trackers

How they work

- Based on small devices which mesure the acceleration which they have been submitted.
- They detect acceleration (accelerometres) and the rotatory orientation (giroscops).
- By integring over the time the acceleration we obtain the velocity; by integring the velocity we obtain the position.





Innertial trackers

Advantages:

 Unlimited area of action. Do not require cables and are not tied to a fixed RS.

Inconvenients:

- The position is relative.
- The errors are accumulative.
- Non exact to mesure absolute position.
- Sensitives to the movement velocity!

How they work:

- Based on image processing.
- Different sort of:
 - Estereo cameras (2 or more known cameras)
 - Làsser
 - Detecting siluets
- Two main techniques:
 - Outside-In
 - Inside-Out

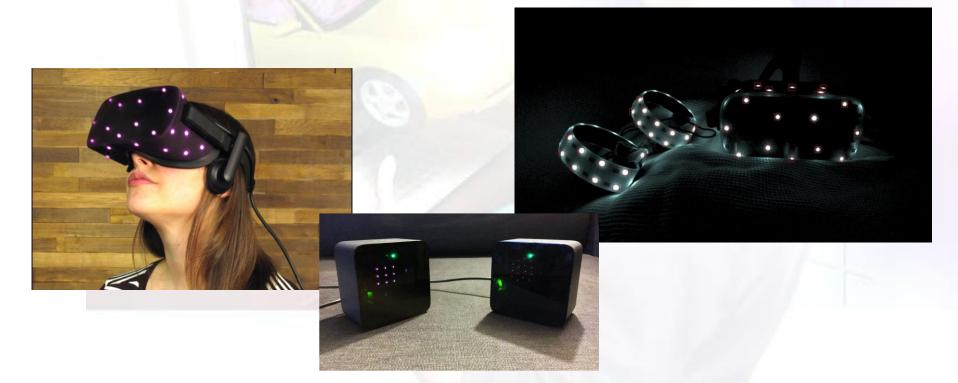


Estereo cameras (2 or more known cameras)

- For each point p to be registered:
 - It detects the position in each image (x, y), (x', y') (correspondence problem).
 - A point (x,y,z) is obtained by intersecting 2 straight lines.
- Require markers or detecting siluetes:
 - Pasive markers: color points
 - Active markers: LEDs

Outside-In:

- Cameras in known locations
 - More cameras improve detection



Inside-Out:

- Cameras on the tracked device looking outward
 - SLAM process generates 3D map of the environment



General advantages

- They can register a big quantity of points
- Relatively cheap

General disadvantages

- Line-of-sight
- Overhead (or dedicated processor).
- Latency time

Outside-In

- √ More accurate readings
- √ Lower latency than Inside-Out
- x Limited work space to the FOV of cameras

Inside-Out

- √ Larger play spaces
- √ Adaptable to new environments
- x More processing time and latency

- Oculus Rift
 - Viewer
 - Resolution: 2160 x 1200
 - FOV: 110



- Oculus Rift
 - Tracking system
 - Accelerometer + giroscope + infrared optical tracking system 360 degrees
 - Area: 1,5 x 3,3







- Oculus Rift
 - Input devices
 - Position + Orientation of commander
 - Detection of finger movements (only 3 gestures)



- HTC-Vive
 - Viewer

Resolution: 2160 x 1200

• FOV: 110



- HTC-Vive
 - Tracking system
 - Accelerometer + giroscope + double laser tracking system
 - Area: 4,5 x 4,5
 - Base stations emit light through stationary LEDs plus a pair of spinning lasers.
 - LEDs flash at 60Hz, and one of the two spinning lasers sweep a beam of light across the room.





- HTC-Vive
 - Tracking system
 - Accelerometer + giroscope + double laser tracking system
 - Area: 4,5 x 4,5
 - The receiver is covered with photosensors to detect flashes and laser beams.
 - The headset starts counting until its photosensors gets hit by a laser beam.
 - This info is used to calculate position relative to base stations



- HTC-Vive
 - Input devices
 - Position + Orientation of commander
 - NO Detection of finger movements





- Valve Index
 - Viewer
 - Resolution: 2880 x 1600
 - FOV: 130

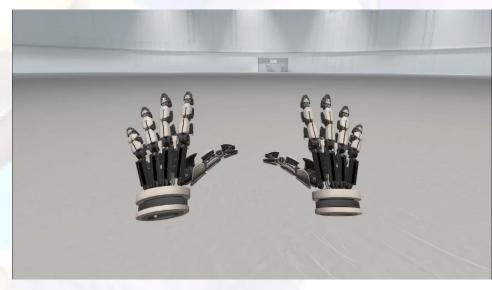


- Valve Index
 - Tracking system (like Vive and improved)
 - Accelerometer + giroscope + double laser tracking system
 - Area: 6.5 x 6,5 (2) 10 x 10 (4)



- Valve Index
 - Input devices
 - Position + Orientation of commander
 - Detection of finger movements





Data gloves











Data gloves

Function

- Allow to detect the position of the hand fingers, expresed by the flexion angles of each finger.
- Three angles per finger (except the thumb 2).
- They are used combined with trackers.

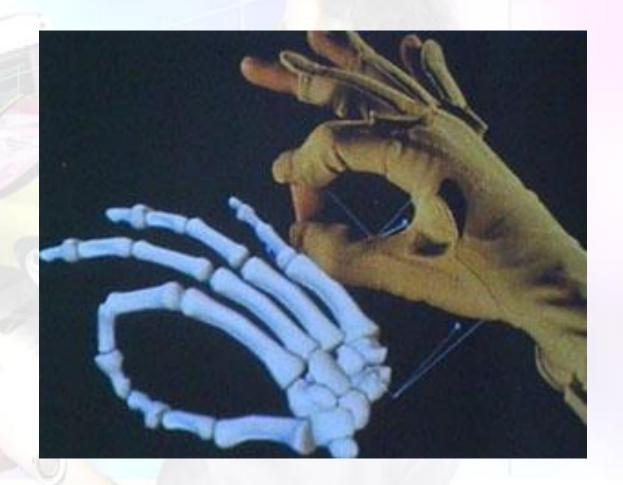
Aplications

- Direct manipulation of objects.
- Intuitive interaction (navigation, menus)
- Interaction based in signals

Data gloves

Technologies

- Optical fiber
- Mechanical
- Optical



Gloves of optical fiber

How they work

- Atenuation of the light which goes through the optical fiber.
- Each finger: one or two circuits of optical fiber especialy treated.
- The attenuation of light depends on the curvature.
- Diference between emmitted light and received light.



Gloves of optical fiber

Advantages

- Very precise
- Ergonomic
- Veri fast

Inconvenients:

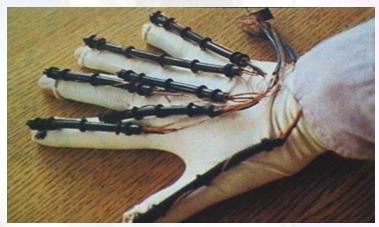
- Require constant calibrations.
- The curvature of the fiber do not always correspond to the one on the fingers (friction).



Mechanical gloves

How they work

- Similar to mechanical trackers.
- The potenciometres are incorporated in a structure similar to an exoesquelet.
- Other models mesure the elongation of a telescopic tube.





Mechanical gloves

Advantages

- Exacts and precise
- Very fast
- Possibility of force-feedback.

Inconvenients

- Expensive
- Non ergonomic







Optical gloves

How they work:

Based on LEDs or passive markers, captured by video

LEDs d'infraroig

cameras.

Advantages:

- No cables.
- Ergonomic.

Inconvenients:

- Line-of-sight
- Lack of precision. High latency.



Microsoft Kinect

- New controller for Microsoft's Xbox 360
- Full-body tracking, face and voice recognition
- Low cost (99 Euros as of April 2011)





Main components



Main components

Video

- Color CMOS camera
- Infrared (IR) CMOS camera
- Infrared projector 830nm, 60mW laser diode.

Audio

4 microphones

Tilt control

- Motor
- Accelerometer (3-axes)

Processors & memory

- PrimeSense chip PS1080-A2
- 64 MB DDR2 SDRAM

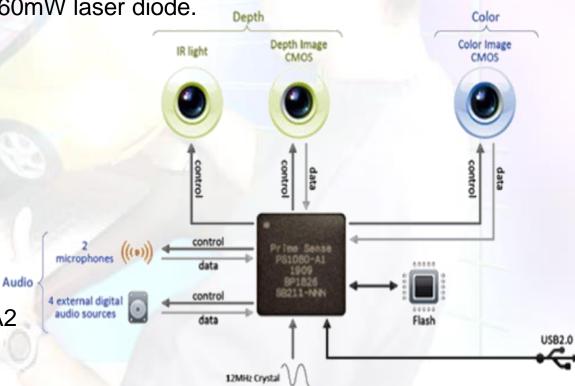


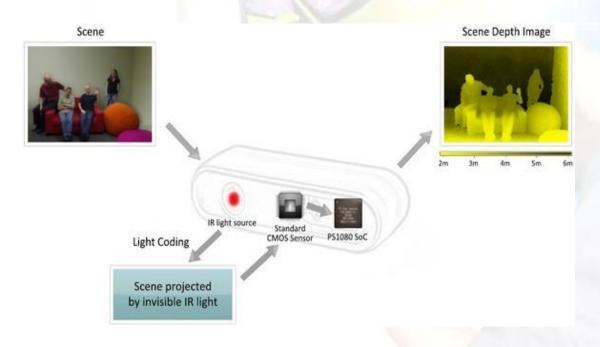
Image sensors

- Color camera: 640x480 sensor, 640x480@30fps output
- IR camera: 1280x1024 sensor, 640x480@30fps output
- Operation range (depth sensor) = 0.4m 3.5m
- FOV = 58° H, 45° V, 70° D
- Spatial resolution (@ 2m distance) = 3mm
- Depth resolution (@ 2m distance) = 1cm

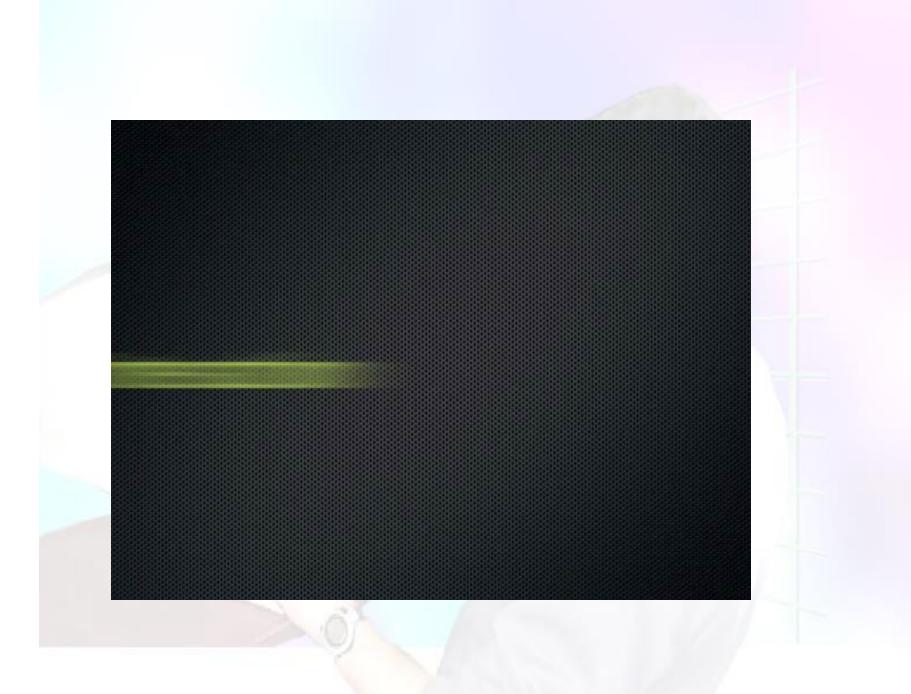


Depth sensing

- The IR emitter projects an irregular pattern of IR dots of varying intensities.
- The IR camera reconstructs a depth image by recognizing the distortion in this pattern.









Overview

- 1. Output devices
 - a) Visual display characteristics
 - b) Visual display systems
- 2. Input devices
 - a) Concepts and classification
 - b) Obtaining position (trackers)
 - HMDs + tracking
 - c) Data Gloves
 - d) Kinect
- 3. Stand-alone systems
- 4. Devices ad-hoc

Stand-alone systems

- Vive Focus Plus (just an example)
 - Viewer
 - Resolution: 2880 x 1600
 - FOV: 110
 - CPU & Memory: Snapdragon 835, 32GB ROM, 4GB RAM



Stand-alone systems

- Vive Focus Plus
 - Tracking system
 - HMD: SLAM (Inside-Out)
 - Accelerometer + giroscope +
 - Simultaneous Location And Mapping (SLAM) algorithm tracking system (Inside-Out using computer vision)
 - Area: --
 - Controllers:
 - 6DoF Ultrasonic + IMU fusion tracking





Overview

- 1. Output devices
 - a) Visual display characteristics
 - b) Visual display systems
- 2. Input devices
 - a) Concepts and classification
 - b) Obtaining position (trackers)
 - HMDs + tracking
 - c) Data Gloves
 - d) Kinect
- 3. Stand-alone systems
- 4. Devices ad-hoc

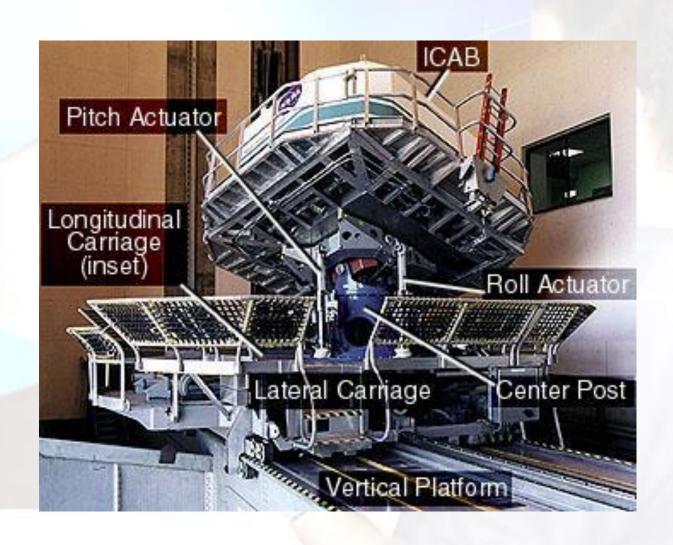
Car simulators

Car simulators

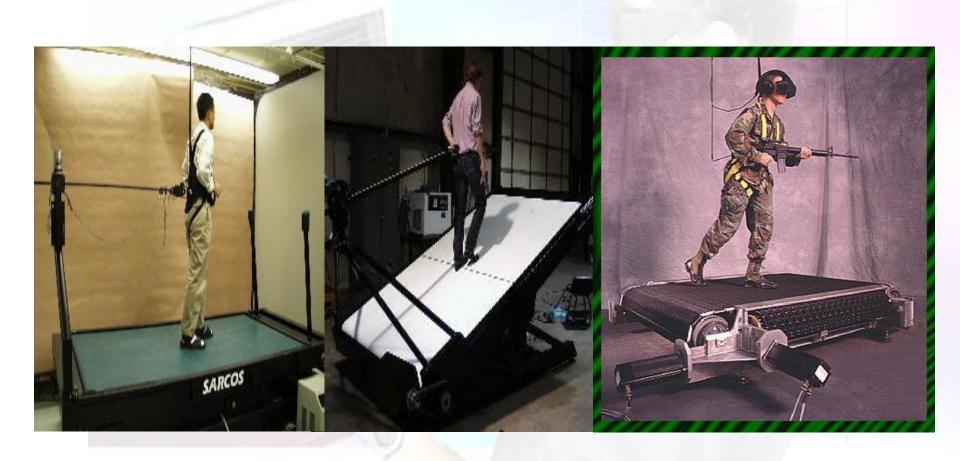




Flight simulators

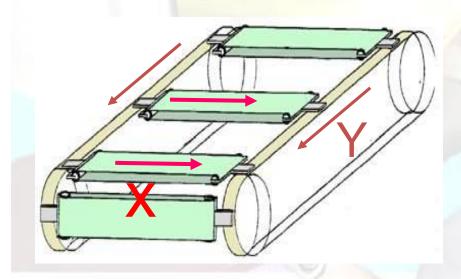


Locomotion interfaces



Omni-Directional treadmill

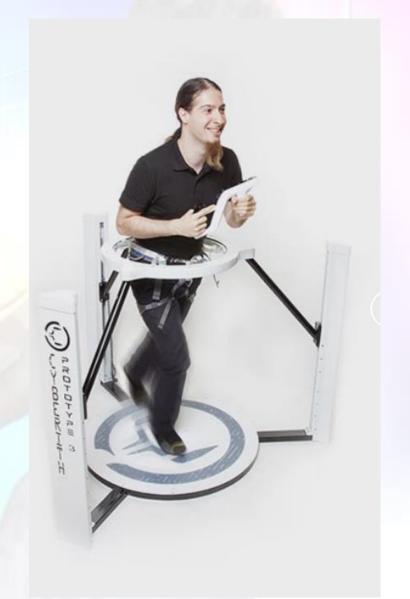
- Surface moves on any direction
- 1.3 m x 1.3m
- Max speed: 3 m/s





Omni-Directional treadmill





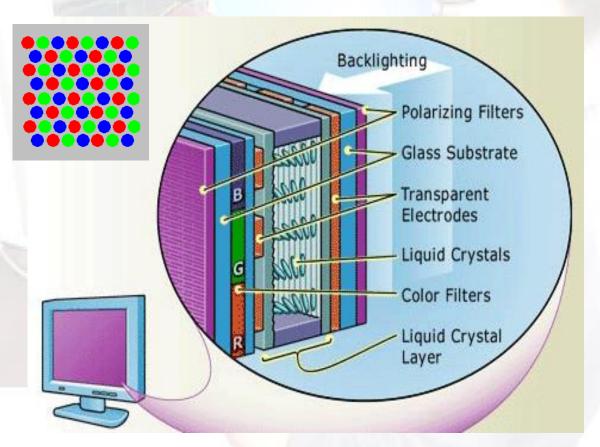
VIRTUAL REALITY

Input and output devices.

Display technologies (extra slides)

Technologies: LCD

Use 3 transmissive panels (one per color) that modulate the light through polarization



Technologies: LCD

Pros:

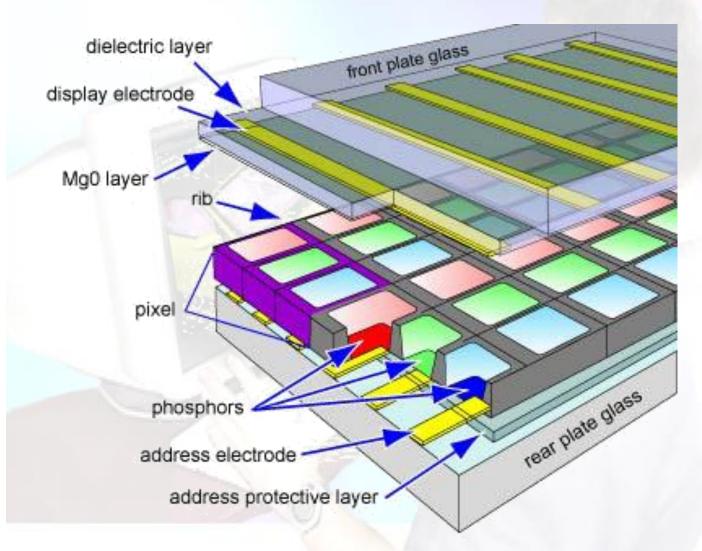
- High brightness levels
- Compact, lightweight
- The light is already polarized

Cons:

- Visual angle
- Best image quality only at native resolution
- Limited contrast levels,
- So-called screen door effect, caused by the transistors and signal wires running between the pixels (the fine lines separating the projector's pixels become visible)
- Response time (eg 6-16 ms typically; refresh rate typically limited to 60Hz)



Technologies: PDP



Technologies: PDP

Pros:

- Each pixel is addressed individually
- Image quality
- Visual field
- Thin

Cons:

- Price
- Memory effect
- Best image quality only at native resolution



Technologies: CRT projectors

Configurations

- A single color CRT
- A single B/W CRT with color wheel
- Three separate CRTs (R, G, B)

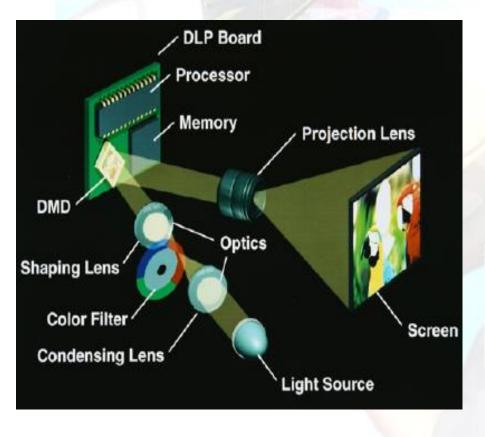


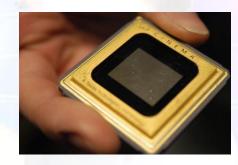
Technologies: LCD projectors

- LCD projectors are like LCD panels
- Backlight is replaced by a powerful lamp
- LCD panel modulates color (like a slide projector)

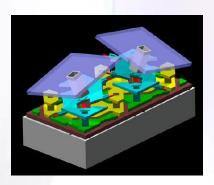


Based on a Texas Instruments chip with millions of micromirrors

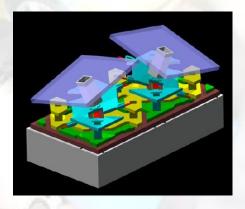






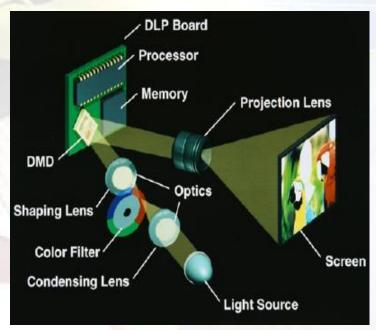


- Each micromirror can be deflected either in the "on" or the "off" position.
- Grey levels are obtained by pulse width modulation, vibrating the micromirrors at high frequencies.



Two main versions:

- Single-chip DLP: creates a full-color image through sequential display of R, G, B images using a color wheel
- Three DLP chips process the R G B images in parallel.



• Pros:

- Compact size (especially for single-chip),
- High brightness levels and good contrast,
- Good picture quality and stability.
- 3-chip DLP projectors are capable of the high refresh rates (96-120 Hz) to display flicker-free active stereo.
- Now available at full-HD

Cons:

Single-chip DLPs suffer from rainbow effects

VIRTUAL REALITY

Input and output devices.

About orientation (extra slides)

Direction cosines

The orientation is represented by 9 values:

```
u_x, u_y, u_z << Unit vector in the direction of X v_x, v_y, v_z << Unit vector in the direction of Y v_x, v_y, v_z << Unit vector in the direction of Z
```

- The rotation matrix is obtained by having the 3 vectors in columns.
- Is the more complete and intuitive representation.
- Is very redundant (9 reals!)

Euler angles

- The orientation is represented by 3 values:
 - x, y, z << Rotation angles with respect each axe
- The rotation is obtained by:

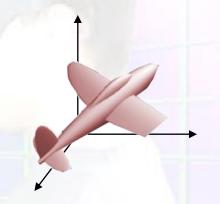
Rot_x, Rot_y, Rot_z

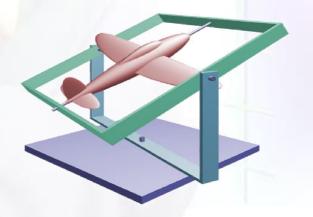
- In order to avoid the ambiguation the following should be fixed:
 - an order to apply the rotations (XYZ, XZY...)
 - an interpretation (alibi, alias).
- Is very intuitive.

Euler angles (ii)

Two interpretations or ways to visualize a rotations composition R_x , R_y , R_z :

- Alibi: All rotations are done with respect to one of the three axes of a Global Reference System which is fixed, in this order R_x, R_y, R_z.
- Alias: The rotations are done with respect to a Local Reference System attached to the object, in the reverse order: R_z, R_y, R_{x.}





Euler angles (iii)

Problems:

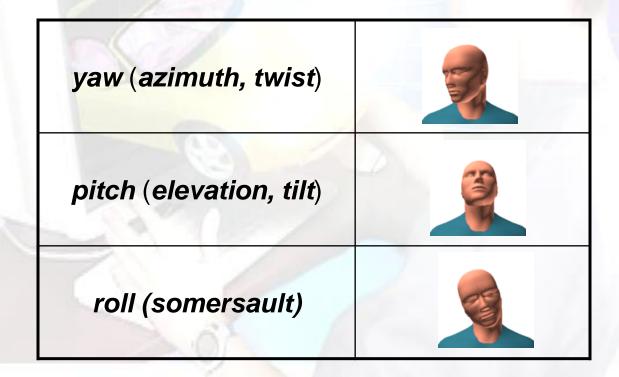
– The rotations are not defined in a unique form:

$$(z, x, y) = (90, 45, 45) = (45, 0, -45)$$

 The cartesian coordinates are independent among them; the Euler angles no.

Euler angles (v)

 Sometimes, some mnemotechnics are used for the angles, which come from the aeronautic:



Angle axis representation

The orientation is represented with four values:

```
x, y, z << Unit vector (arbitrary rotation axe) angle << Rotation angle with respect to the vector
```

 The rotation is obtained by rotating angle grades around the axe determined by the vector (x, y, z).

Spherical coordinates

The orientation is represented by 3 values:

```
long, lat << Longitude/latitude of a unit vector angle << Rotation angle with respect to the vector
```

 Is equivalent to the angle axis representation, but the unit vector is represented in spherical coordinates.

Quaternion

The orientation is represented with four values:

```
x, y, z << Unit vector (arbitrary rotation axe)
w << Rotation with respect to the vector
```

- The rotation is obtained by rotating 2*acos(w) around the axe determined by the vector (x, y, z).
- Is a notation very similar to the angle axis representation; it has advantages in animation, because it's easy to interpolate between two orientations.
- Is less intuitive.