XR HARDWARE

Dr Samit Bhattacharya

Computer Science and Engineering

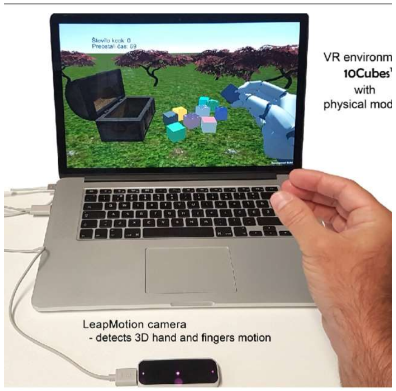
IIT Guwahati

HARDWARE DEVICES

- Input devices
- Computer
- Output devices

INPUT DEVICES

- Devices through which user provides input to XR system
- Send signals to system about user action [1]



TYPES

- Hand input devices
- Non-hand input devices

HAND INPUT DEVICES

- Devices that requires hand to operate

- FIVE types

TYPE (1) - WORLD-GROUNDED INPUT DEVICES

- Devices fixed in real world (e.g. joystick)

TYPE (1) - WORLD-GROUNDED INPUT DEVICES

- Most often used to interact with desktop systems
  - Constrains user movement

- Good for VR vehicle metaphor
  - Used in location-based entertainment

Fig: The Disney Aladdin world-grounded input device along with its mapping for viewpoint control. (From Pausch et al. [1996])

TYPE (2) - NON-TRACKED HANDHELD CONTROLLERS

- Held in hand, but not tracked in 3D
  - Buttons
  - Joysticks
  - Triggers
  - Traditional video game controllers (e.g., Xbox controller)

TYPE (2) - NON-TRACKED HANDHELD CONTROLLERS

- Issue: break-in-presence
  - User moves hands but sees no corresponding movement in virtual hands from its current position
  - Due to non-tracking

TYPE (3) - TRACKED HANDHELD CONTROLLERS

- Physical device with 6 DOF tracking inside (e.g. Vive controllers)
  - Combines joystick input with tracking
  - Majorly used in VR

TYPE (3) - TRACKED HANDHELD CONTROLLERS

- Natural, direct mapping to hand motion
  - Viewpoint manipulation typically using buttons and a trackball
- Cons - user not able to directly/fully touch and feel other passive objects

TYPE (4) – HAND-WORN DEVICES

- Devices worn on hands/arms
  - Gloves
  - EMG bands
  - Rings
  - …

DATA GLOVES

- Bend sensing gloves
  - Passive input device
  - Detects hand posture and gestures
  - Continuous raw data from bend sensors
    - Fiber optic, resistive ink, strain-gauge
  - Large DOF output, natural hand output

DATA GLOVES

- Pinch gloves
  - Conductive material at fingertips
  - Determine if fingertips touching
  - Used for discrete input
    - Object selection, mode switching, etc.

TYPE (4) – HAND-WORN DEVICES

- Advantages
  - Natural input with potentially rich gesture interaction
  - Hands can be held in comfortable positions – no line of sight issues
  - Hands and fingers can fully interact with real objects

TYPE (4) – HAND-WORN DEVICES

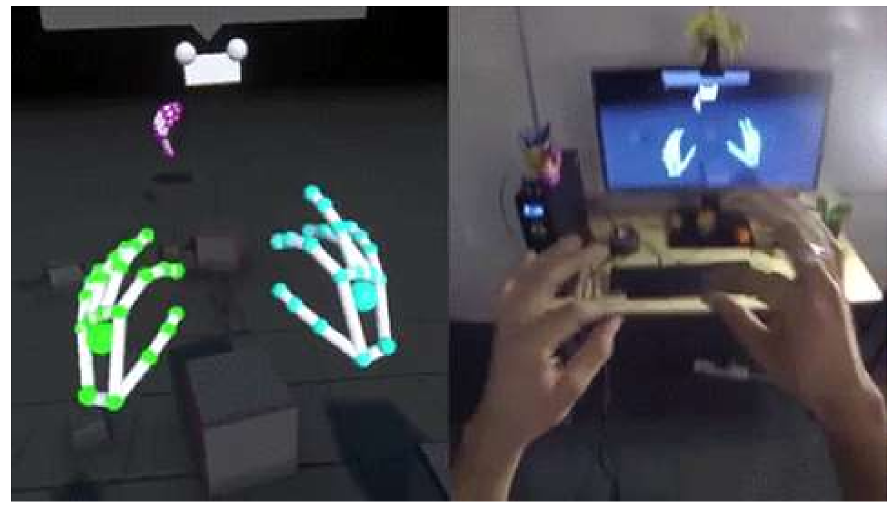
- Issues
  - Recognizing more than a few gestures requires the user to recalibrate often
  - Uncomfortable – result in sweaty hands

TYPE (5) – BARE HAND INPUT DEVICES

- Tracks bare hand input (computer vision)

- Natural interaction

- User's hands are completely unencumbered



EXAMPLE - LEAP MOTION

- IR based sensor for hand tracking
  - HMD + Leap Motion = Hand input in VR
- Equipped with 3 IR LEDs and 2 wide angle cameras
- LEDs generate IR lights, picked up by the camera after getting reflected with the hands
- Hand tracking done by software
- 1m range, 0.7 mm accuracy, 200 Hz

TYPE (5) – BARE HAND INPUT DEVICES

- Challenging
  - Does not provide for ‘sense of touch’ – reduces natural-ness
  - Fatigue from holding hands in front of sensor
  - Line-of-sight requirements
  - Consistent recognition of gestures across a wide range of users

NON-HAND INPUT DEVICES

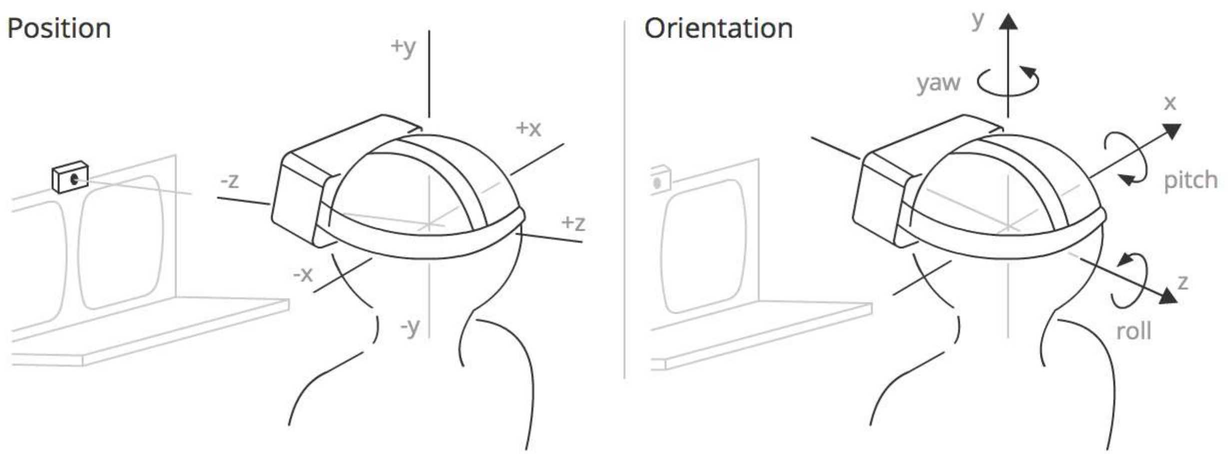
- Captures input from other parts of the body
  - Head movement
  - Eye gaze
  - Voice
  - Body movement
  - …

NON-HAND INPUT DEVICES

- FOUR types

TYPE (1) - HEAD TRACKING

- Use head motion for input



Source: https://media.prod.mdn.mozit.cloud/attachments/2015/06/04/11083/afb780cabae3c25ffdc10618f68a628c/positionOrientationVR.png

TYPE (1) - HEAD TRACKING

- Most common form of head-tracking interaction is to aim by looking
  - Provide a reticle or pointer in the middle of the screen that is triggered by a button press

TYPE (2) - EYE TRACKING

- Tracks where the eyes are looking
- Effective for specialized tasks and subtle interactions
  - How a character responds when looked at

Eye tracker integrated into VR HMD

TYPE (2) - EYE TRACKING

- The Midas Touch problem
  - People expect to look at things without that look "meaning" something
  - Interacting via eye tracking alone is usually not a good idea - works better with multimodal input

TYPE (3) – FULL BODY TRACKING

- Tracking more than just the head and hands
- Adds to the illusion of self-embodiment as well as social presence

TYPE (3) – FULL BODY TRACKING

- Approaches
  - Using a motion capture suit (MoCap) – various sensors within the suit
  - Using camera based tracking
    - Microsoft Kinect
    - OptiTrack

MOCAP SUITS

The Perception Neuron suit – contains a network of straps with inertial measurement unit (IMU) sensors

https://www.vfxvoice.com/what-mocap-suit-suits-you/

CAMERA BASED MOTION CAPTURE

- Multiple cameras
- Reflective markers on body
- Eg – Optitrack (www.optitrack.com)

TYPE (4) – VOICE/SPEECH/SOUND TRACKING

- **Captures voice, speech, non-speech sounds as input for interaction**
  - **Microphones**
  - **Microphone arrays**

COMPUTER

- Device that process data from input and render on output device
- Selected according to requirement [1]
  - Depend on application, user, I/O devices and level of immersion
  - Could be a standard PC with more processing power and a powerful graphics accelerator
  - Or distributed computer systems interconnected through high speed communication network
  - Or can even be our smartphone!

IMPORTANT SYSTEM COMPONENTS

- GPU
- CPU
- RAM

GPU (GRAPHICS PROCESSING UNIT)

- Data-parallel, throughput oriented processors
- Multi-GPU enabled system used for building VR applications
- Major companies manufacturing GPU: Nvidia, AMD, Intel & Asus

CPU

- Multi-core processors greatly increase computational capacity and are quickly becoming standard

- Virtual environment can be developed once and scale efficiently with number of cores [2]

RAM

- Each XR engine has its own minimum RAM requirements.
- Most VR engines requires at least 8GB of RAM to build VR apps

OUTPUT DEVICES

- Visual displays
- Audio output
- Haptics
- Motion Platforms
- Treadmills
- Other sensory outputs – taste, smell (not widely used/popular)

VISUAL DISPLAY

- THREE broad types
  - Head-mounted displays
  - World-fixed displays
  - Regular displays

DISPLAY TYPE (1) - HEAD-MOUNTED DISPLAY (HMD)

- Most common metaphor for VR

- A sophisticated equipment
  - Integration of electronic, optical, mechanical and even audio components

DISPLAY TYPE (1) - HEAD-MOUNTED DISPLAY (HMD)

- Purpose - to place images directly in front of user's eyes using one (for monoscopic viewing) or two (for stereoscopic viewing) small screens

HMD TYPES

- Non-see-through (used in VR)
- Optical-see-through (used in AR)
- Video-see-through (used in augmented virtuality)

DISPLAY TYPE (2) – WORLD-FIXED DISPLAY

- Render graphics onto surfaces that do not move with head

- Displays range from standard monitor to displays that completely surround the user (CAVE)

WORLD-FIXED DISPLAY: FISH TANK VR

- Introduced by Colin Ware et al (1993) – "…a stereo image of a 3D scene viewed on a monitor using a perspective projection coupled to the head position of the observer"
- Contains [4]
  - CRT monitor
  - Haptic device
  - Optical tracker

SURROUND SCREEN DISPLAY (CAVE)

- Has 3 to 6 large screens
- Puts user in
  a room for visual immersion
- Usually driven by a single or group of
  powerful graphics engines
- Requires elaborate tracking

Source: https://www.evl.uic.edu/pape/CAVE/DLP/dcp_2634.jpg

DISPLAY TYPE (3) – REGULAR DISPLAY

- Refers to any regular display (smartphone, tablet, computer monitor) – not anything specialized
- Also called Information Browsers (mainly used for AR/MR)

https://www.augmented-minds.com/wp-content/uploads/2018/01/augmented-reality-ar-browser.jpg

AUDIO OUTPUT

- Spatialized audio provides a sense of where sounds are coming from in 3D space
- Speakers can be fixed in space or move with the head
- Headphones are preferred for a fully immersive system

HAPTICS

- Output exploiting our touch perception
- Several categorization
  - Passive (static physical objects) or Active (physical feedback controlled by computer)
  - Tactile (through skin) or Proprioceptive force (through joints/muscles)
  - Self-grounded (worn) or World-grounded (attached to real world)

PASSIVE VS ACTIVE HAPTICS

- Passive haptics
  - Real-world physical object matching shape of virtual object
- Active haptics
  - Controlled by a computer - most common form of haptics

TACTILE VS. PROPRIOCEPTIVE FORCE

- Tactile haptics - provide sense of touch through the skin
  - Vibrotactile stimulation - evokes tactile sensations using mechanical vibration of the skin
  - Electrotactile stimulation - evokes tactile sensation via an electrode passing current through the skin
- Proprioceptive force
  - Provides a sense of limb movement and muscular resistance

SELF-GROUNDED VS. WORLD-GROUNDED

- Self-grounded haptics
  - Worn/held by and move with user
  - Hand-held controllers are also examples of self-grounded haptics

- World-grounded haptics
  - Physically attached to the real world and can provide a true sense of fully solid objects that don't move

Source: https://csdl-images.computer.org/trans/th/2017/04/figures/pacch1-2689006.gif

MOTION PLATFORMS

- Hardware device that moves entire body resulting in a sense of physical motion and gravity
  - Convey sense of orientation, vibration, acceleration, and jerks
  - Motion sickness can be reduced by decreasing the conflict between visual motion and motion felt

MOTION PLATFORMS - TYPES

- Active motion platform - controlled by computer simulation

Source: https://www.vrwayvr.com/data/upload/ueditor/20191018/5da9621b42a5c.png

MOTION PLATFORMS - TYPES

- Passive motion platform - controlled by user

Source: https://www.ippinka.com/wp-content/uploads/2017/08/ICAROS-virtual-fitness-09.png

TREADMILLS

- Provide a sense that one is walking or running while actually staying in one place

Fig: The Virtuix Omni: Passive omnidirect
ional treadmill

TREADMILLS

- Omnidirectional treadmills
  - **Active** - have computer- controlled mechanically moving parts
  - **Passive** - contains no computer controlled mechanically moving parts

TRACKING AND TRACKERS

TRACKERS

- Special-purpose hardware to *capture/detect* real-time change in object’s position and/or orientation

WHAT TO TRACK

- Amount of translation/rotation/both in
  - User's sense organs: Head, Eyes
  - Other body parts: Limbs, Fingers
  - Entire body
  - Rest of environment: movable objects
  - Other people in physical space

WHY TO TRACK

- Many reasons
  - View control
  - Locomotion
  - Object manipulation
  - Controlling an *avatar* (virtual representation of user)

HOW TO TRACK

- Complex process
- Things to consider
  - Tracker position
  - Tracking process
  - Tracker types

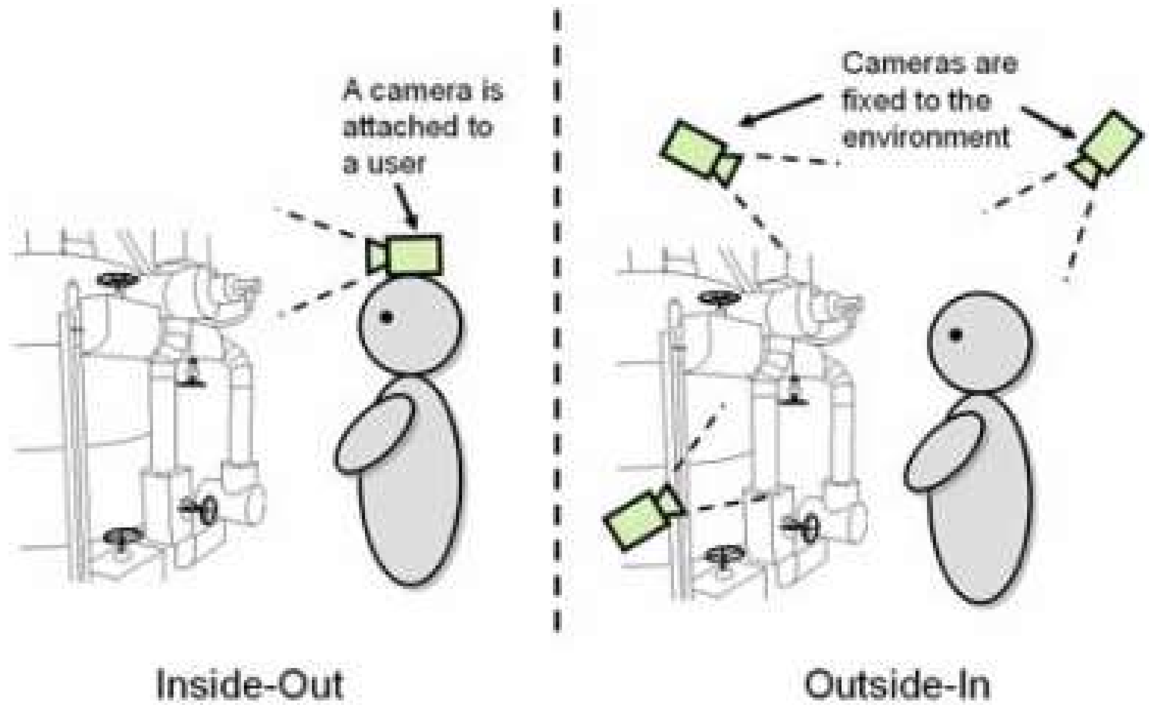
OUTSIDE-IN TRACKING

- External sensors, cameras, or markers required
  - Tracking constrained to specific area
  - Used by most VR headsets today

INSIDE-OUT TRACKING

- Camera or sensor located on HMD
  - No need for other external devices to do tracking
  - Examples
    - Microsoft HoloLens
    - Intel Project Alloy
    - Qualcomm VR820
- Uses Simultaneous Localization And Mapping (SLAM) technique

INSIDE-OUT V/S OUTSIDE-IN TRACKING



[Source: https://xinreality.com/mediawiki/images/5/5a/Inside_out_vs._outside_in_tracking.png](https://xinreality.com/mediawiki/images/5/5a/Inside_out_vs._outside_in_tracking.png)

TRACKER TYPES

- Electromechanical
- Electromagnetic
- Acoustic/Ultrasonic
- Optical
- Inertial
- Others

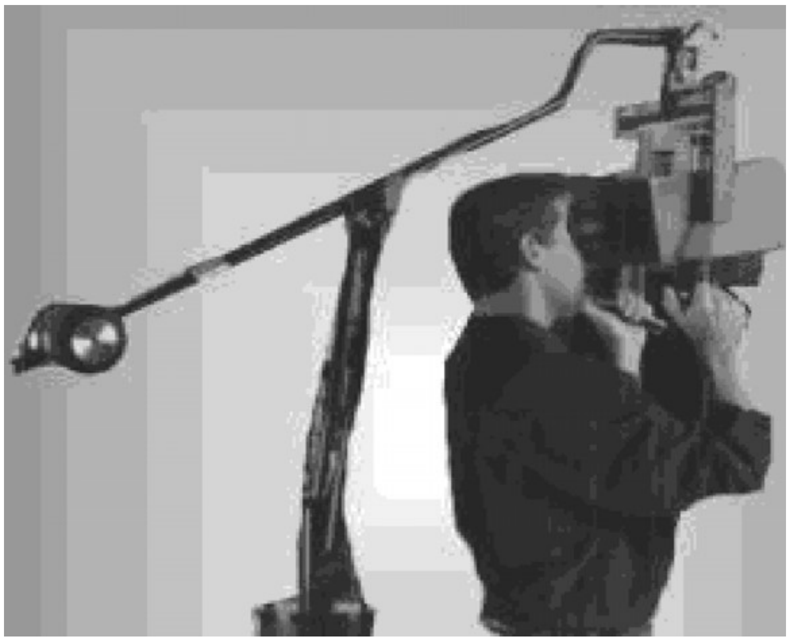
TYPE (1) - ELECTROMECHANICAL TRACKERS

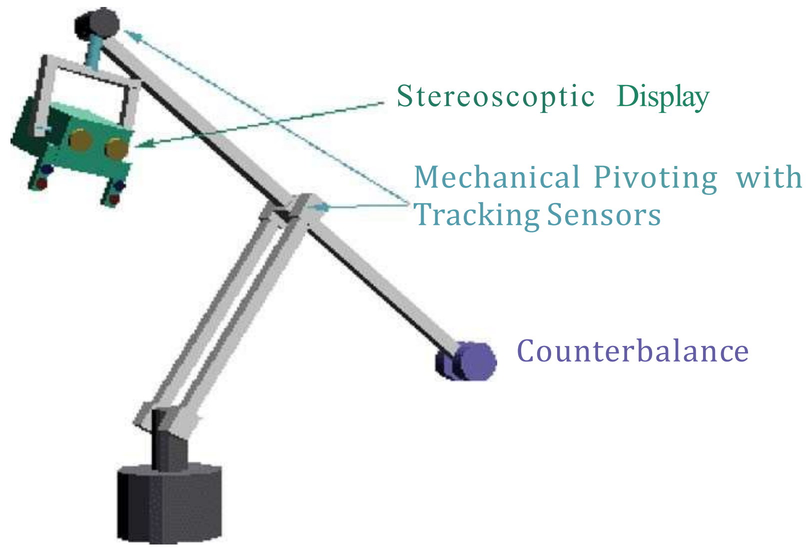
- Consists of a serial or parallel kinematic structure composed of links interconnected using joints with sensors

  - Can be world-fixed or worn on the body

BOOM

Display attached to electromechanically tracked booms





Fake Space Boom

ANIMAZOO

- Animated avatars
- Mechanical tracker
- Several products
  - IGS-190
  - Gypsy-6
  - GypsyGyro-18
  - Exoskeleton tracker

http://www.animazoo.com/

TYPE (2) - ELECTROMAGNETIC TRACKERS

- A non-contact position measurement device

- Uses a magnetic field produced by a *stationary transmitter* to determine real-time position of a *moving receiver* element

TYPE (3) - ACOUSTIC TRACKING

- Uses ultrasonic signal produced by stationary transmitter to determine real-time position
- Transmitter (speakers), microphones (receivers)
  - Orientation: several transmitters
- Requires line-of-sight

TYPE (4) - OPTICAL TRACKERS

- Uses optical sensing to determine real-time position/orientation
- Works through **triangulation**
- Sensing component
  - CCD or CMOS camera
  - Photodiode
  - Photosensor

WORKING – CAMERA TRACKING

- Outside-in - sensing component(s) fixed
  - Light sensors or cameras surrounding tracked object

- Some markers or light beacons typically placed on user
  - Passive reflectors/fiducial markers - not always robust identification of markers
  - Active IR-transmitters (LED) - robust identification, but requires some wires

WORKING - VIDEOMETRIC TRACKING

- Inside-looking-out - video camera on tracked object
    - Camera watches surroundings

- System analyzes images
    - Try to locate landmarks
    - Derive camera's relative position w.r.t landmarks

- Single camera can determine own 6-DOF position
    - Multiple reference points needed

TYPE (5) - INERTIAL TRACKERS

- Self-contained sensors that measure rate of change in object orientation, translation velocity or acceleration
  - Only changes, not absolute value
- Microelectro-mechanical systems (MEMS) technology
- No range limitations
- Integrated into most smartphones

TYPE (5) INERTIAL TRACKERS

- Gyroscopes – measures rate of change in angular velocity

- Accelerometers – measures acceleration

- Magnetometer – measures magnetic field change

- ….

OTHERS - MOTION CAPTURE

- Technique of digitally recording movements of real things (usually humans)
- Can be based on any tracking method
- Often 1-2 facial and 4-5 full body cameras used
  - Often not real time

OTHERS - EYE TRACKING

- Head-mounted or remote
- Optical tracking – hi-res camera-based
- Infrared lighting, electro-occulography (EOG), video oculography (VOG) …
- Tracks gaze point/area

OTHERS - HYBRID TRACKERS

- Utilizes two or more position/orientation measurement technologies

- Tracks objects better than any single technology alone would allow

TRACKING STEPS

- FOUR steps

TRACKING STEPS

- **Calibration**: sensor output transformed to behave as closely to real value as possible

TRACKING STEPS

- **Integration**: aggregation/integration of sensor measurements at discrete points in time

TRACKING STEPS

- **Registration**: estimation/determination of initial sensor position and orientation

TRACKING STEPS

- **Drift correction**: compensating for difference between estimated position and orientation and the actual value (drift error)

PROCESS EXPLANATION (BRIEF)

- Make use of IMU (Inertial Measurement Unit) – a composite hardware containing three types of sensors (part of HMD)
  - **Gyroscope** – to capture angular velocity
  - **Accelerometer** – to capture liner acceleration (along X, Y, Z axes)
  - **Magnetometer** – to capture magnetic field strength along 3 axes

PROCESS EXPLANATION (BRIEF)

- Tracking relies mainly on gyroscope readings (angular velocity)
- Other sensors (accelerometer & magnetometer) useful for detecting drift errors

- Technical issues easy to understand in 2D case – we'll do that
  - Same concepts extend to 3D tracking

2D TRACKING

- Measured angular velocity of gyroscope (ωg) and the actual angular velocity (ωa) likely to be different – due to imperfect device and measurement
  - **Calibration error:** $\omega g - \omega a$
- Note: we can relate the two values as follows

  $\omega g = a + b.\omega a$ (if both are same, a=0, b=1)

- Thus, calibration error: $\omega g - \omega a = a + b.\omega a - \omega a = a + (b-1).\omega a$

CALIBRATION

- Involves taking many samples (only once, initially), sometimes thousands
    - From cheap and inaccurate sensors (ω_g)
    - And also expensive and highly accurate sensors (ω_a)

- Compare those values (simple approach: compute sum of squares error)

$$\sum_{1}^{n}(\omega_g - \omega_a)^2$$

CALIBRATION

- Select constants c1 and c2 to optimize

$$\sum_{1}^{n} \left(c_1 + c_2\omega_g - \omega_a\right)^2$$

  - A classical regression problem referred to as linear least-squares
- Once c1, c2 determined, cheap sensor readings can be calibrated as

$$\omega\text{cal} = \text{c1} + \text{c2}.\omega\text{g}$$

- Essentially maps cheap sensor output to expensive sensor, without actually using one!

INTEGRATION

- Sensor outputs usually arrive at a regular sampling rate
- Orientation θ(t) at time t (=k.Δt, Δt = sampling interval) can be computed by *integrating* the samples

$$\theta(t) = \theta(0) + \sum_{i=1\ to\ t}(\omega_{cal}^{i}.\Delta t)$$

- **Idea:** calibrated angular velocity sample at i-th instant (ω_{cal}^{i}) indicates rotation (i.e., change in orientation) by an amount ($\omega_{cal}^{i}.\Delta t$)

REGISTRATION

- Setting initial alignment between real and virtual worlds [$\theta(t=0)$]
  - Whichever direction headset is facing when tracking system turned on
  - Setting initial alignment to be as *forward direction*
- Fiducial markers, reflectors, or tags aid registration significantly

REGISTRATION

- Visual, sound, haptics registration
- Indoors/outdoors
- Static / dynamic error

DRIFT CORRECTION

- First thing – detect drift error
- Usually done with the help of extra sensors (e.g., camera)
- A classis approach to blend two sensors: gyroscope & camera

$$\theta_d(k) = \alpha.\theta_g(k) + (1- \alpha).\theta_c(k)$$

$\theta_d(k)$ = drift corrected orientation at time instant t

$\theta_g(k)$ = orientation obtained by integrating gyroscope values at k-th instant

$\theta c(k)$ = orientation obtained with camera at k-th instant

$0 \leq \alpha \leq 1$ called “gain parameter”

THANK YOU

REFERENCES

- Jerald, J., 2015. The VR book: Human-centered design for virtual reality. Morgan & Claypool.
- Steven M. LaValle, Virtual Reality, http://lavalle.pl/vr/
- Björn Blissing., 2016., A technical report on tracking techniques for automotive virtual reality., VTI notat, 25A-2016., *Swedish National Road and Transport Research Institute*

REFERENCES

[1] Oluleke Bamodu, X. Ye, 2013. Virtual Reality and Virtual Reality System Components. Proceedings of the 2nd International Conference On Systems Engineering and Modeling (ICSEM 2013) pp.921-924.

[2] Acosta, Eric & Liu, Alan & Sieck, Jennifer & Muniz, Gilbert & Bowyer, Mark & Armonda, Rocco, 2009. A Multi-core CPU Pipeline Architecture for Virtual Environments. Studies in health technology and informatics. pp.10-12.

[3] Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola, Jr. Ivan Poupyrev, 2005. Handbook: 3D User Interfaces: Theory and Practice pp.31-40.

[4] Wen Qi, Russell M. Taylor II, Christopher G. Healey, Jean-Bernard Martens, 2006. A Comparison of Immersive HMD, Fish Tank VR and Fish Tank with Haptics Displays for Volume Visualization. APGV '06: Proceedings of the 3rd symposium on Appliedperception in graphics and visualization. pp.51–58.

REFERENCES

[5] Billinghurst, M. Clark, A. Lee, G. (2015). **A Survey of Augmented Reality.** *Foundations and Trends® in Human-Computer Interaction, 8, pp. 73-272*. Link, pages 165-177

[6] Van Krevelen, R. & Poelman, R. (2010). **A Survey of Augmented Reality Technologies, Applications and Limitations**. *International Journal of Virtual Reality, 2010, 9(2),pp. 1-20* . Link, pages 7-10

[7] Kassahun Bekele, M. Pierdicca, R. Frontoni, E. Savina Malinverni, E. & Gain, J. (2018). **A Survey of Augmented, Virtual, and Mixed Reality for Cultural Heritage**. *J. Comput. Cult. Herit. 11(2), Article 7 (June 2018), pp. 36 pages*. Link, pages 10-11