ISO/IEC JTC 1/SC 24 "Computer graphics, image processing and environmental data representation"
Secretariat: **BSI**
Committee Manager: **Whitlock Charles Dr**

N 4378 Multimodal VTS 2020 07 28_final

Document type	Related content	Document date	Expected action
Meeting / Presentation	Meeting: VIRTUAL Jul-Aug 2020	2020-08-03	**INFO** by 2020-08-03

Multimodal Interfaces in Virtual Training Systems

ISO/IEC JTC 1/SC 24Plenary & WG Meetings

28 July – 17 August 2020

Kwan-Hee Yoo, YoungSik Bae
Chungbuk National University, Whoborn

Use cases

- Vehicle Driving Simulator

https://www.youtube.com/watch?v=uOMmIfXTUNc

3

Use cases

- Ship Flight Simulator

https://www.youtube.com/watch?v=kPKRDfRDLU8

Use cases

- Excavator Simulator

Use cases

- Excavator Simulator

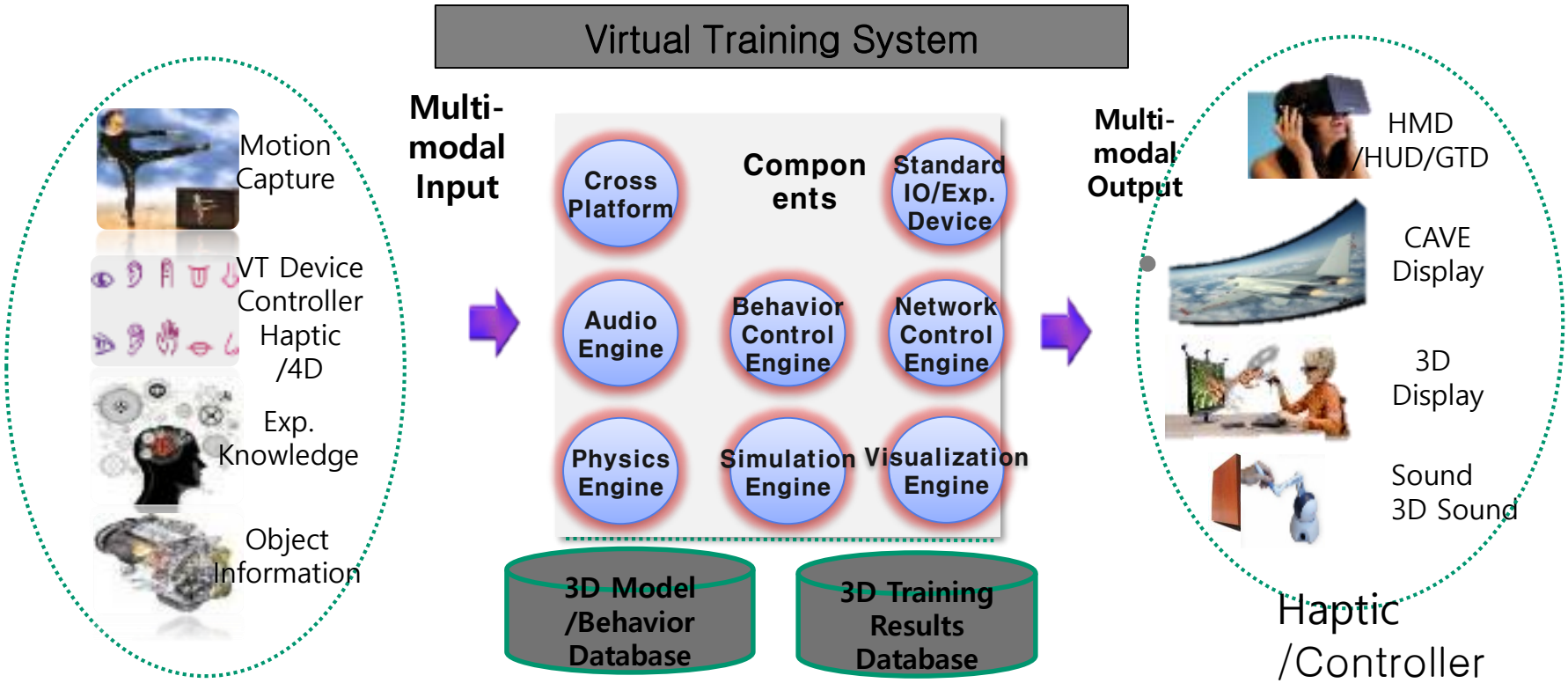
Concepts

Virtual Training System

- A training system that provides a safe and inexpensive education and training more effectively than actual training while realizing immersive experiences by computer simulation under virtual environment similar to actual manufacturing / defense / medical / disaster.

- It is possible to overcome temporal and spatial limitations by minimizing the risk of work environment

Virtual Training System



Components in Virtual Training System

Components	Issues
Instruments	Fitness with virtual training systems
	Safety of Instruments
Instrument Operation	Manufacturing fitness with virtual training system
	Accuracy of operation control program
	Safety of Operating Equipment
Electrical equipment	Manufacturing fitness with virtual training system
	Safety of electrical equipment
Imaging System	Manufacturing fitness with virtual training system
	Performance of Imaging System
Network System	Seamless Training Data Transmission System

Components in Virtual Training System

Components	Issues
Multi-modal display device	Accuracy of multi-modal control
	Program accuracy of multi-modal control
	Safety of multi-modal devices
Content Production	Object Model
	Reproduce object movement
	Reproduce object rendering
Simulation SW	Accuracy of Simulation Interface
	Accuracy of Training Simulation Action
	Safety of Training Simulation Action
Virtual Training Evaluation System	Virtual Training Evaluation Components based on Education Evaluation Tool

SC24 Components in Virtual Training System

Components	Issues	Scope
Visualization	3D rendering Physics Engine Special Effect	CG/VR/AR/Simulation Multi-Modal Interfaces -HMD/GTD/HUD
Behavior Control	Movement/Status of Objects Behavior DB Control Behavior Visualization	CG/VR/AR
Interfaces	Standard I/O devices Experimental HW devices Real-time Data Processing	Multi-Modal Interfaces -Haptic/Controller
Sound	3D Sound	Multi-Modal Interfaces -2D/3D sound
Training evaluation	Training DB Training Items and Evaluation	Education

Properties of multimodal interface in VTS

- Transparency
  - ✓ the condition for the human operator to perform the task in VE without perceiving the effect of presence of the mediating system generating the appropriate stumuli
- Fidelity
  - ✓ the ability of the VE interface system to generate appropriate stimuli without being affected by disturbing effects during time.
- Three Components
  - Vision (Rendering)
  - Audition
  - Haptics/Controller

Properties of multimodal interface in VTS

- Vision (Rendering): A large number of perceptual primitives.
  - ✓ Cortical visual pathways: where(dorsal) stream, what(ventral) stream
  - ✓ What: the analysis of form and color
  - ✓ Where: the analysis of motion and spatial relations
  - ✓ Perception: color, space, motion
  - ✓ HMD(head mounted display)
  - ✓ GTD(glasses type display)

Properties of multimodal interface in VTS

- Audition: the localization of sound in 3D spaces.
  - ✓ auditory analysis of scenes
  - ✓ Music and speech perception
  - ✓ Sound feedback

Properties of multimodal interface in VTS

- Haptics/Controller: exploits tactile and kinesthetic stimuli to acquire information from contact condition of the user with the external environment.
  - ✓ Kinesthetic: the sense of balance, sense of touch
  - ✓ Somato sensation
  - ✓ What stream: perceiving objects and their properties
  - ✓ Where stream: the localization of objects with respect to a user's frame of reference.
  - ✓ Perception-action loop

Applications of multimodal interfaces in Virtual Training Systems

Medical Training System

Medical Training System

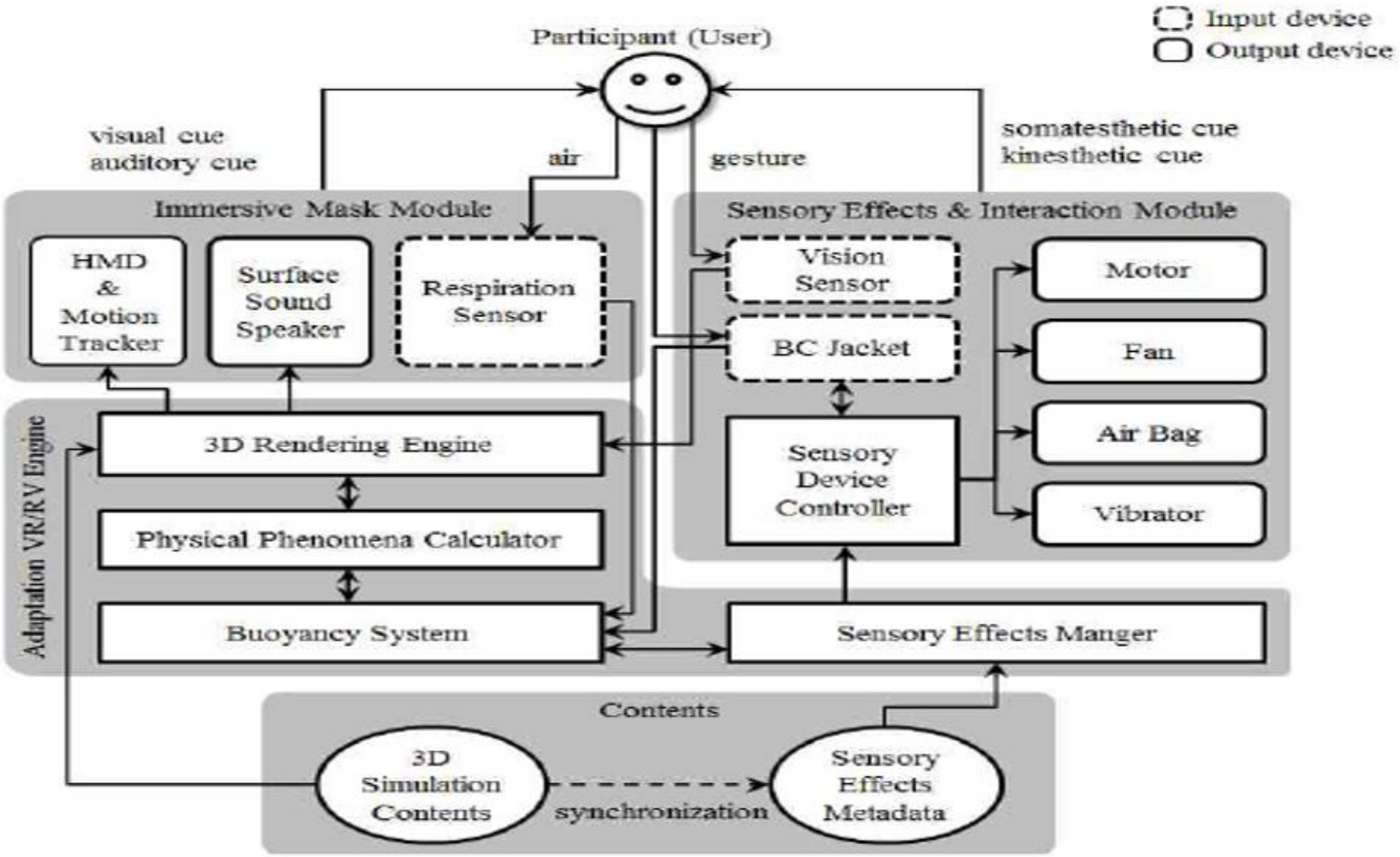
- Vision
  - Projector-based augmented reality system for interventional visualization inside MRI scanners
- Haptics/Controllers
  - Foot-based interaction: Navigation, Sensor Floor
  - Eye+Foot+Hands: multimodal interaction in the operating room
    - ✓visualization: 2D Monitor, HMD
    - ✓foot: navigation
    - ✓eye: panning and zooming
    - ✓hands: surgery instrument control)
- Audio
  - sound: 2D/3D

Virtual Training in VR dome projection system

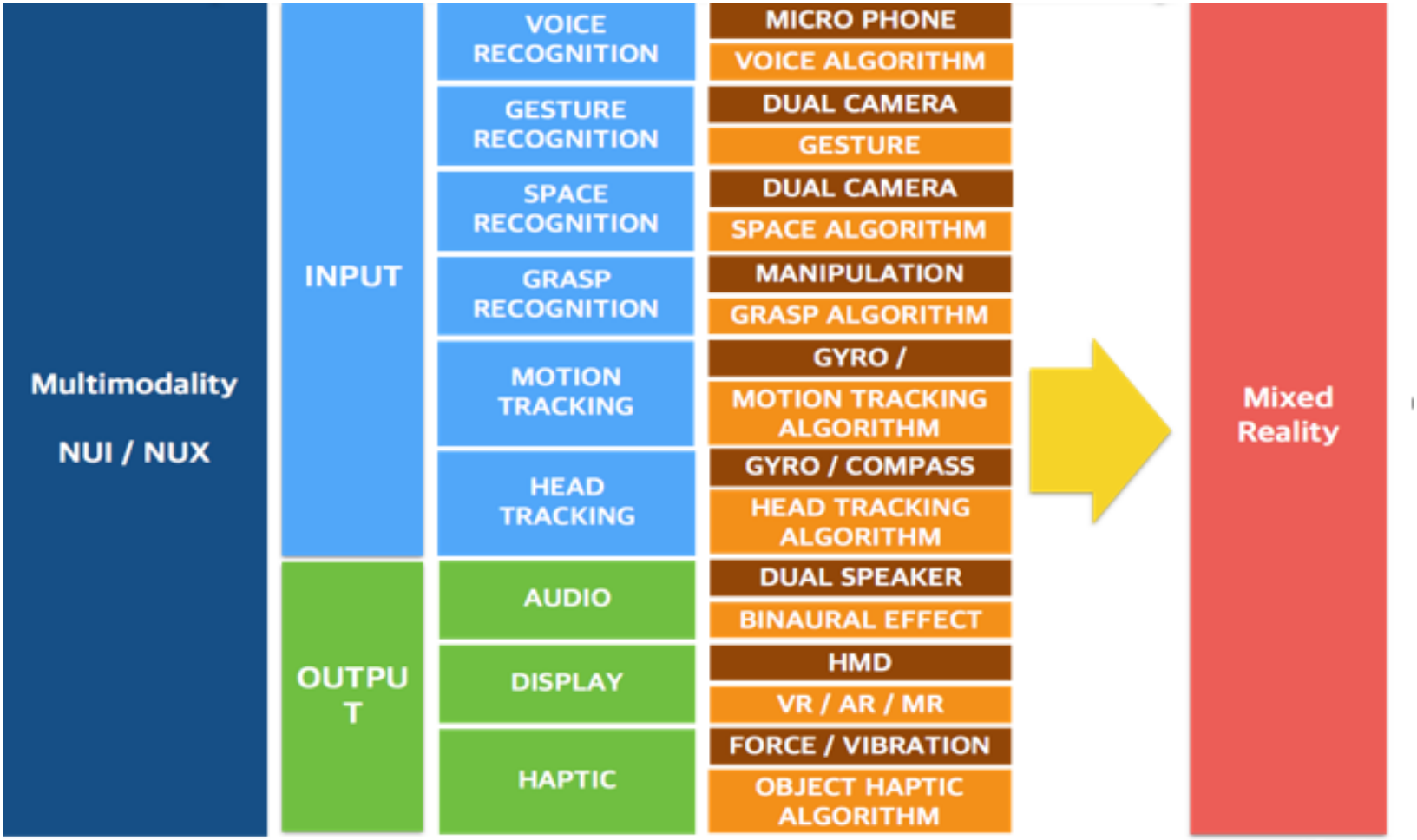
Virtual Training in VR dome projection system

- Vision
  - Multiple projection cameras
- Haptics/Controllers
  - Gestures/Controller (glove)
- Audio
  - sound: 2D/3D

Multimodal Interfaces in Virtual Submarine Operation Simulation



Multimodal Interfaces in Virtual Training System



NUI: Natural User Interface, **NUX**: Natural User Experience, **MR**: Mixed Reality, **VR**: Virtual Reality, **AR**: Augmented Reality

Multimodal Interfaces Format Structure

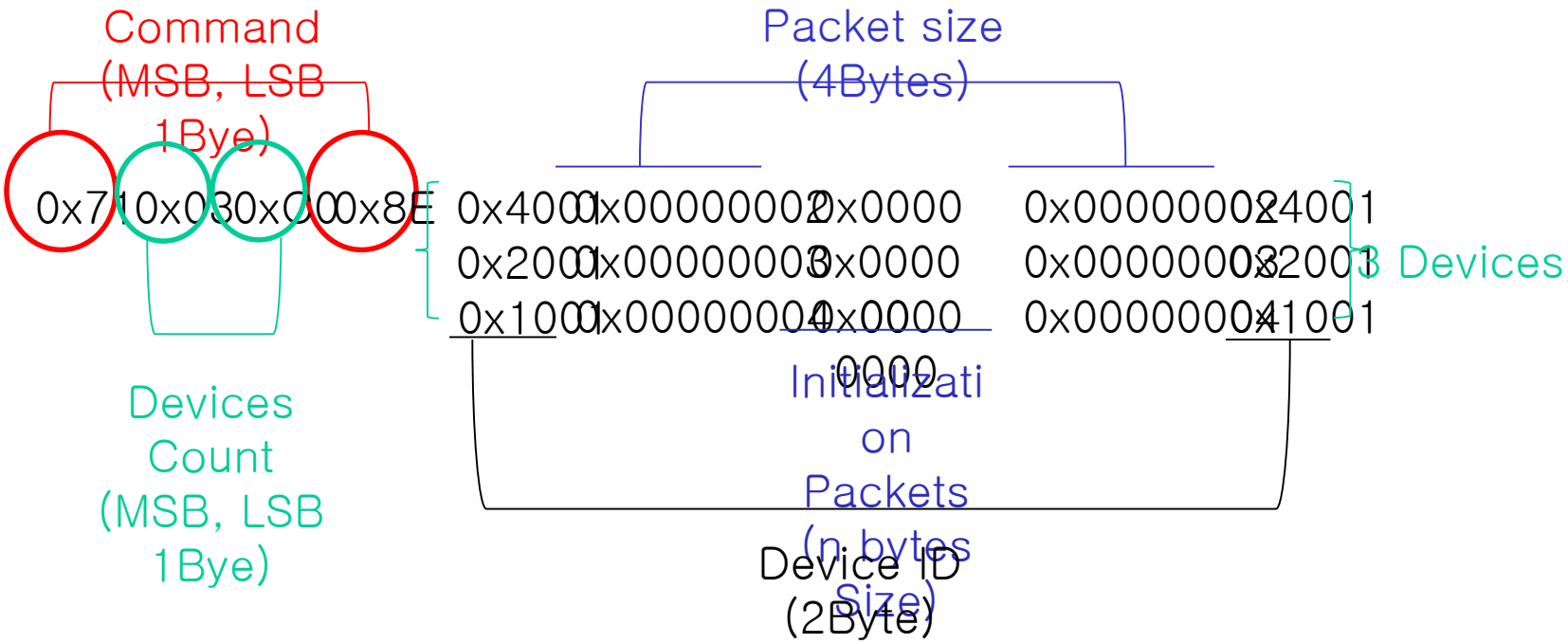
- Multimodality Interfaces Protocol Input/Output Format

I/O	Command (MSB First 1Bye) xSHTSTRx								Devices Count (MSB First 1Bye)	Devices Count (LSB First 1Bye)	Command (LSB 1Bye)							0~n Device ID (2Bytes)	0~n Packet size (4Bytes)	Packets (Packet Size)	0~n Packet size (4Bytes)	0~n Device Number (2Bytes)
	I/O	Sight	Hearing	Touch	Smell	Taste	Reserve	CmdFlag	0x00 ~ 0xFF (max 255)	0x00 ~0xFF (max 255)	CmdFlag	Reserve	Taste	Smell	Touch	Hearing	Sight	(R)SHTSTR(R) + 0x00 ~ 0xFF Device Number	0x00000000 ~ 0xFFFFFFFF byes	Transfer Data Packet	0x00000000 ~ 0xFFFFFFFF byes	(R)SHTSTR(R) + 0x00 ~ 0xFF Device Number
INPUT	1	x	x	x	x	x	x	1	n	revers n	1	x	x	x	x	x	x	0xXX00 ~ 0xXXFF	Command size	Device Input Initialize Command packet Data	Command size	0xXX00 ~ 0xXXFF
		x	x	x	x	x	x	0	n	revers n	0	x	x	x	x	x	x	0xXX00 ~ 0xXXFF	Data Size	Device Input Packet Data	Data Size	0xXX00 ~ 0xXXFF
OUTPUT	0	x	x	x	x	x	x	1	n	revers n	1	x	x	x	x	x	x	0xXX00 ~ 0xXXFF	Command size	Device Output Initialize Command packet Data	Command size	0xXX00 ~ 0xXXFF
		x	x	x	x	x	x	0	n	revers n	0	x	x	x	x	x	x	0xXX00 ~ 0xXXFF	Data Size	Device Output Packet Data	Data Size	0xXX00 ~ 0xXXFF

Multimodal Interfaces Format Example

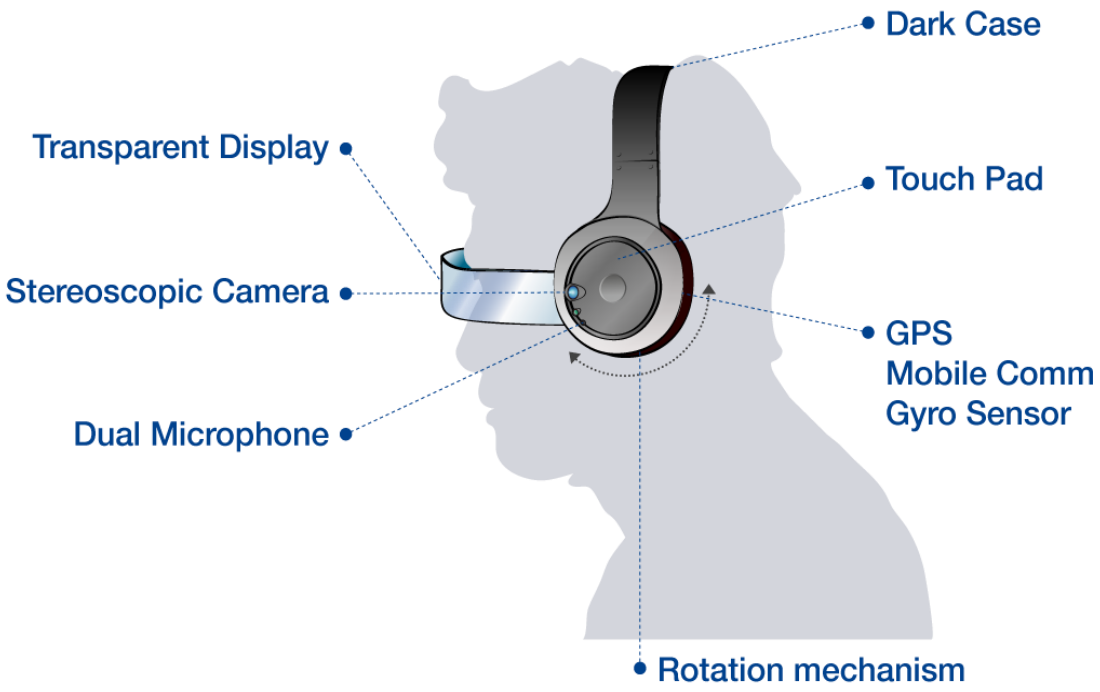
- Output Type 3 Devices Initialization.

Sight(Device ID : 0x4001 – 3D Display) / Hearing(Device ID : 0x2001 – 3D Audio Headphone) / Touch(Device ID : 0x1001 – Haptic Gluves)



Vision Interface Issues

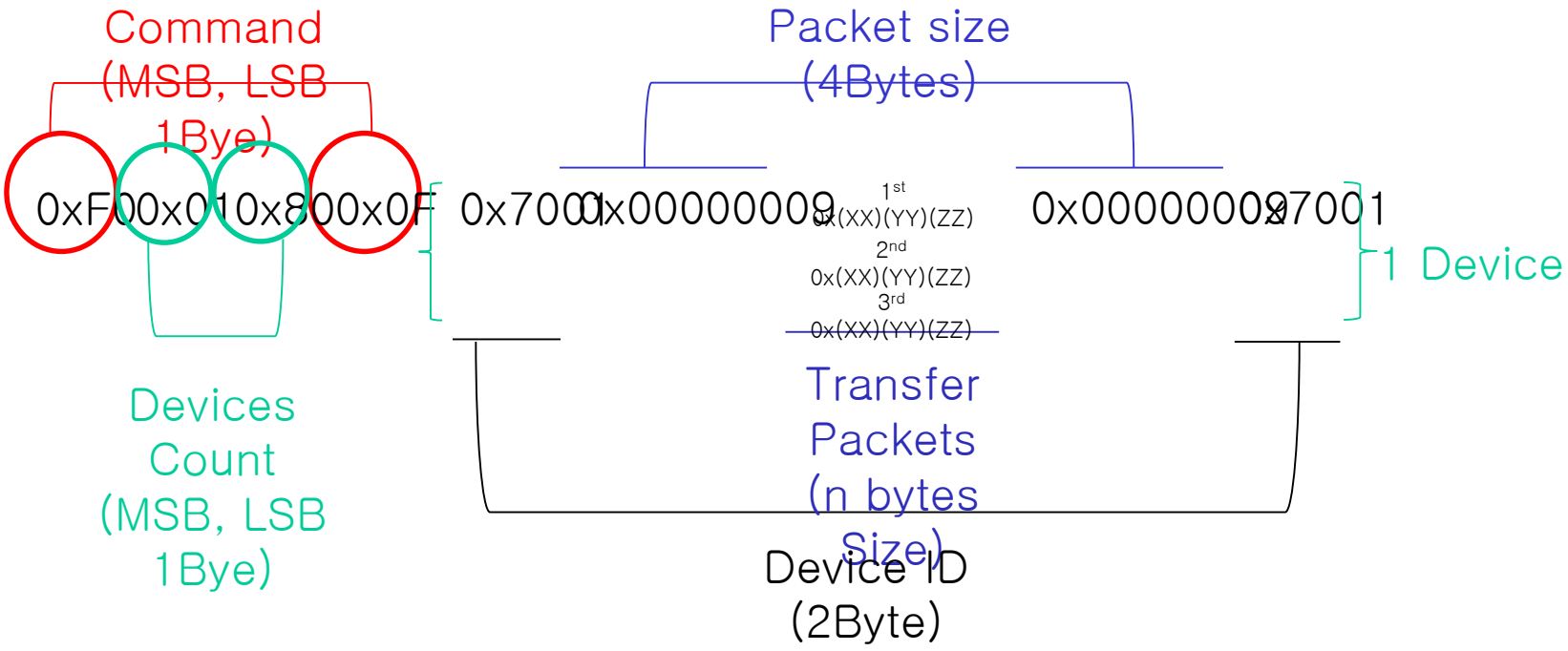
Vision Issues of multimodal Interfaces in Virtual Training



Multimodal Interfaces Input Data Transfer for Vision

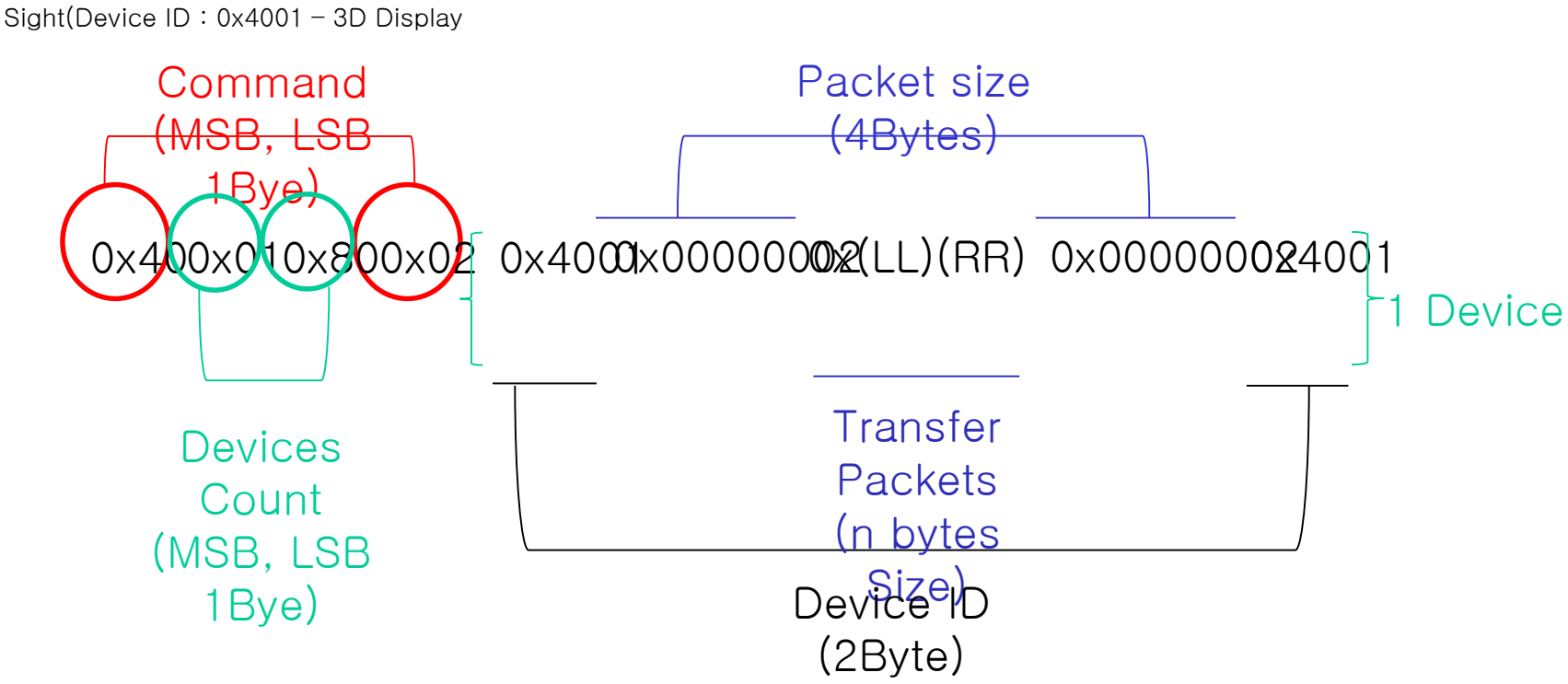
- Input Type 1 Device Data Transfer.

Sight / Hearing / Touch (Device ID : 0x7001 – 9DOF Sensor[Headtracking Sensor])

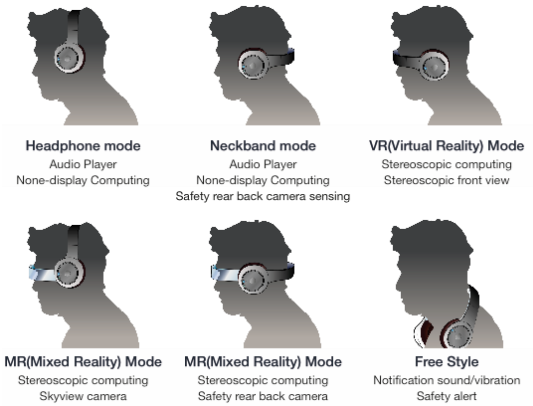


Multimodal Interfaces Output Control Transfer for Vision

- Output Type 1 Device Control Transfer.

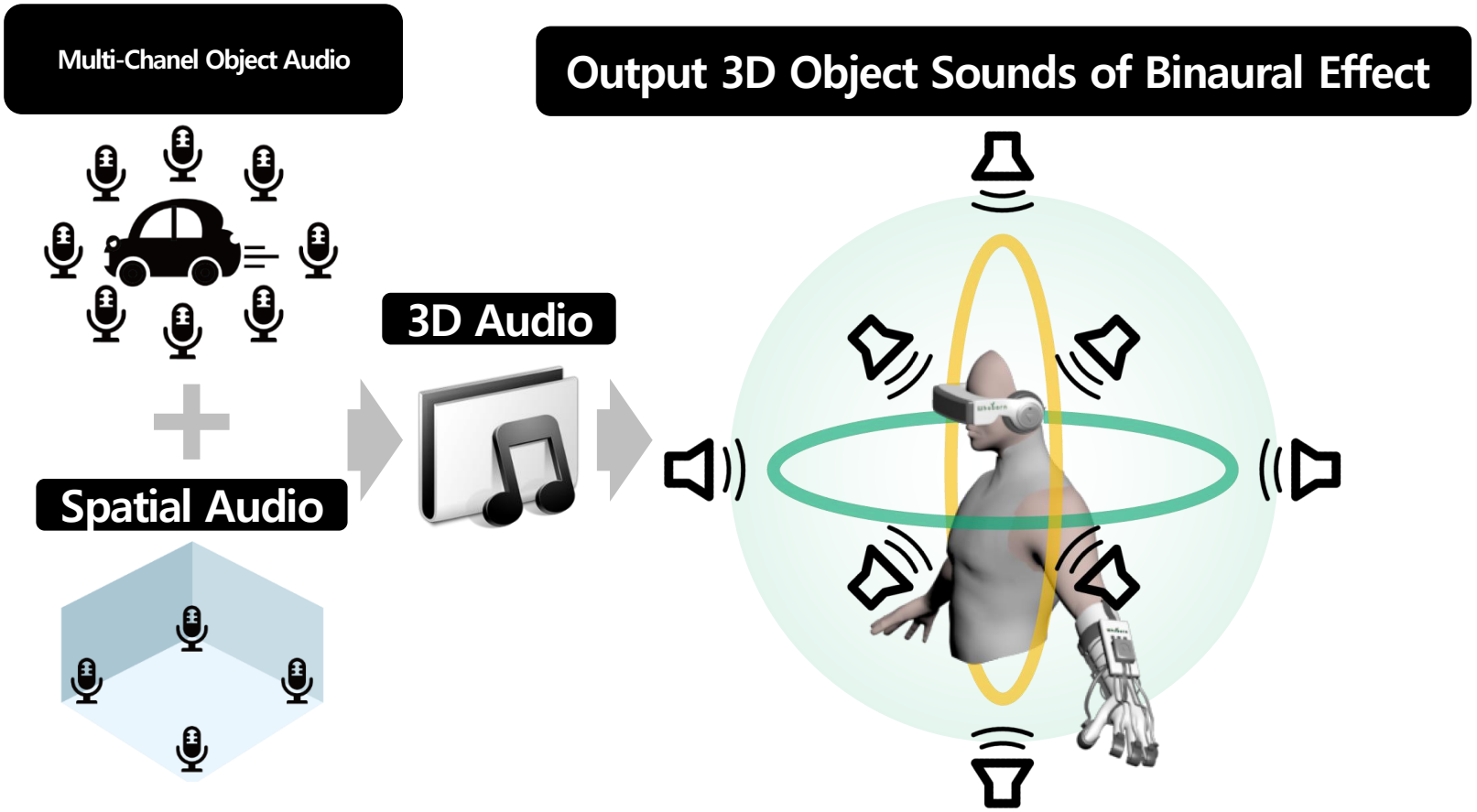


Vision Issues of multimodal Interfaces in Virtual Training



Audition Interface Issues

Audition Issues of multimodal Interfaces in Virtual Training



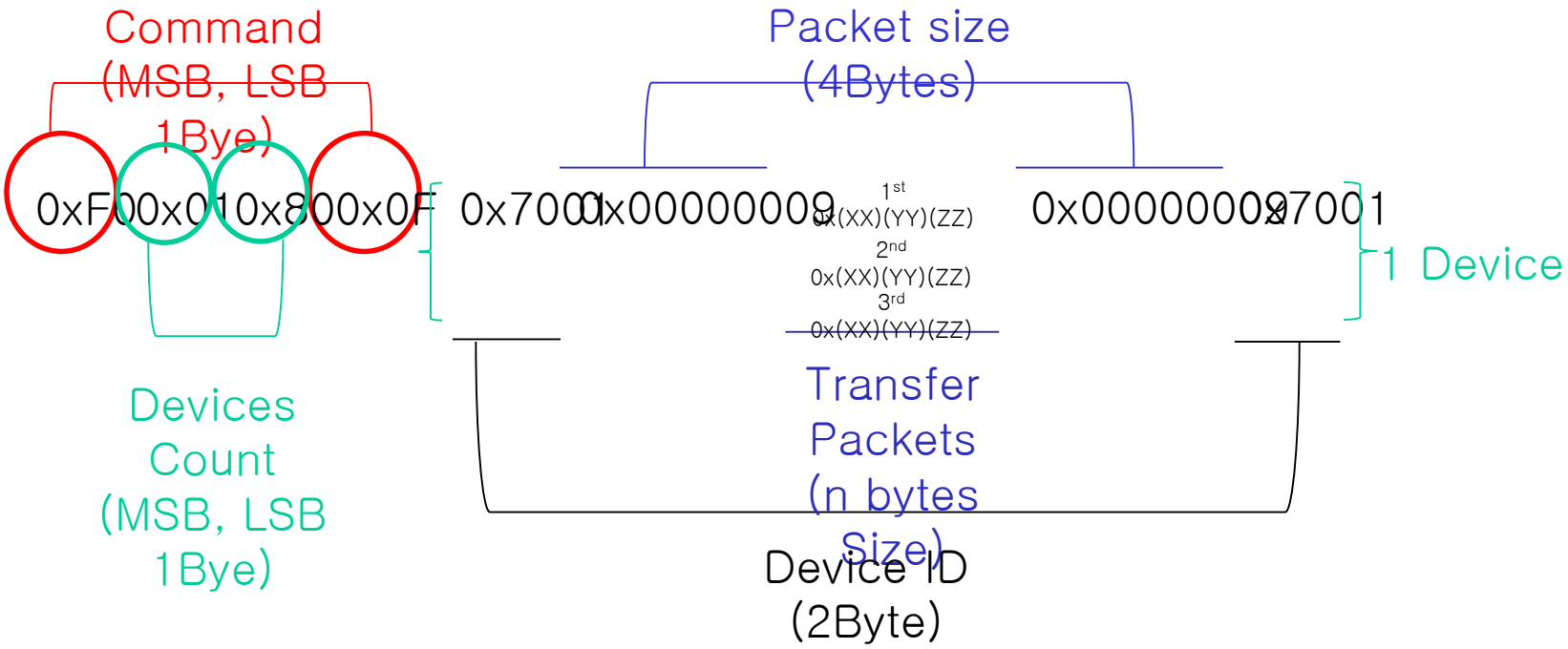
Audition Issues of multimodal Interfaces in Virtual Training

- Multi Microphone (Binaural Recording)
- Spatial Sound
- Headtracking & Move Tracking (9 DOF Sensor : Accelerometer/Gyroscope/Magnetometer)
- Binaural Effect (HRTF Processing)
- Multi Object Sound Muxing

Multimodal Interfaces Input Data Transfer for Audition

- Input Type 1 Device Data Transfer.

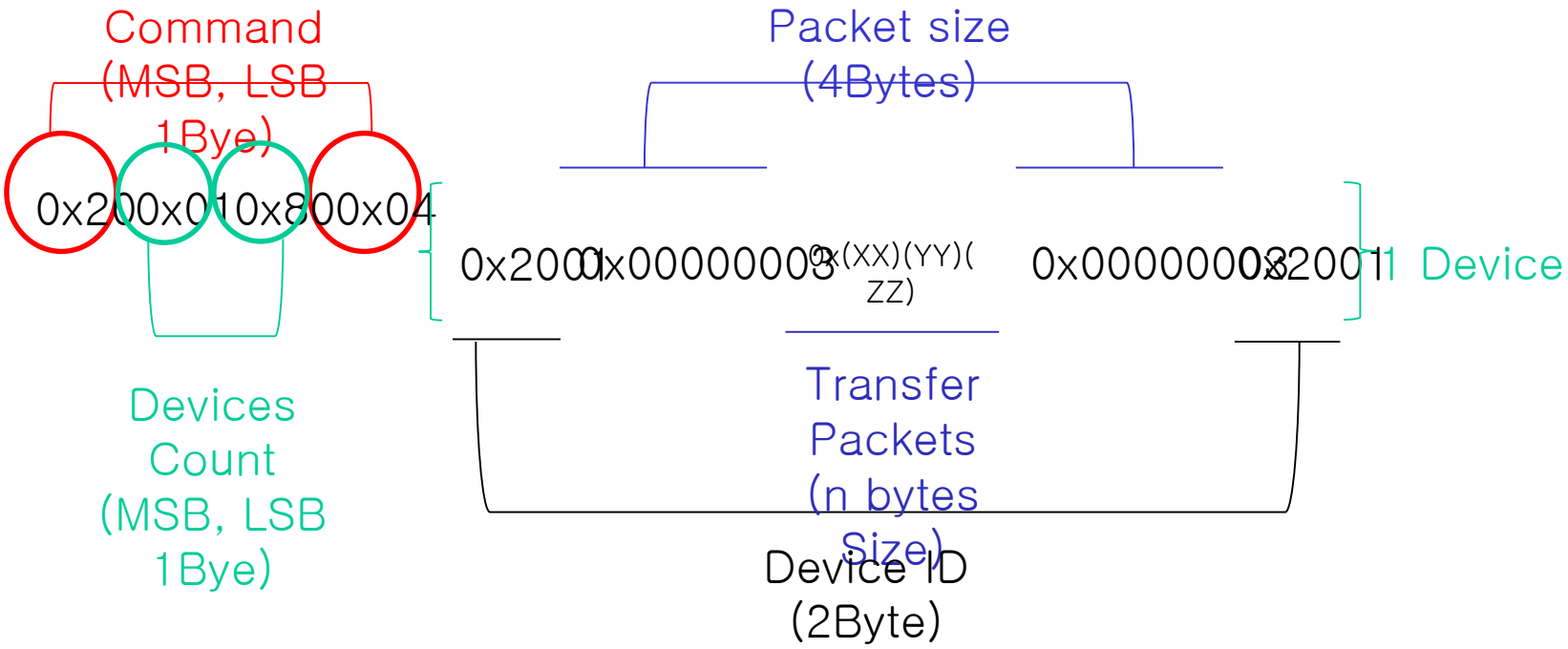
Sight / Hearing / Touch (Device ID : 0x7001 – 9DOF Sensor[Headtracking Sensor])



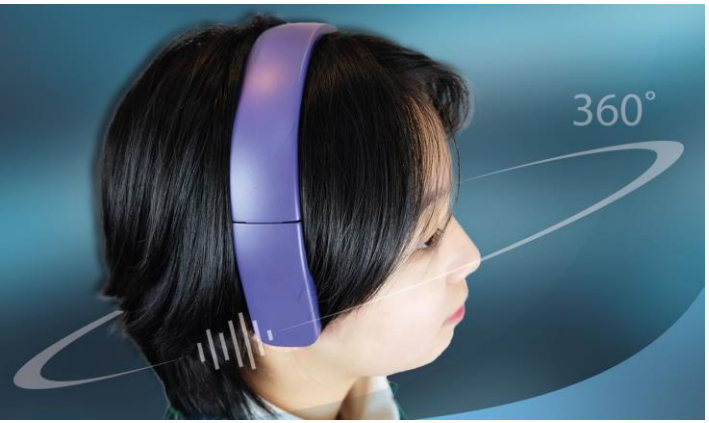
Multimodal Interfaces Output Control Transfer for Audition

- Output Type 1 Device Control Transfer.

Hearing(Device ID : 0x2001 – 3D Audio Headphone)



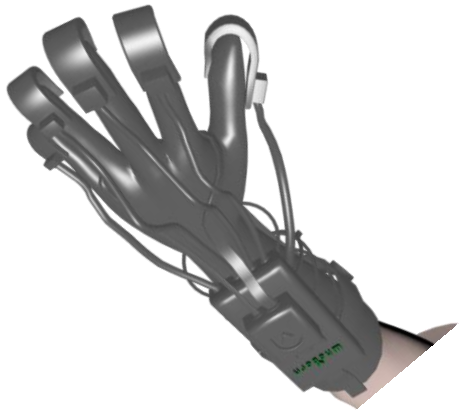
Audition Issues of multimodal Interfaces in Virtual Training



Haptics/Controllers Interface Issues

Haptics Issues of multimodal Interfaces in Virtual Training

- Grasp recognition, Haptic
- Reproduce the tactile sense of the material, pressure of the object in the hand
- Weight due to gravity, intensity of strength, directionality reproduction
- Gesture recognition



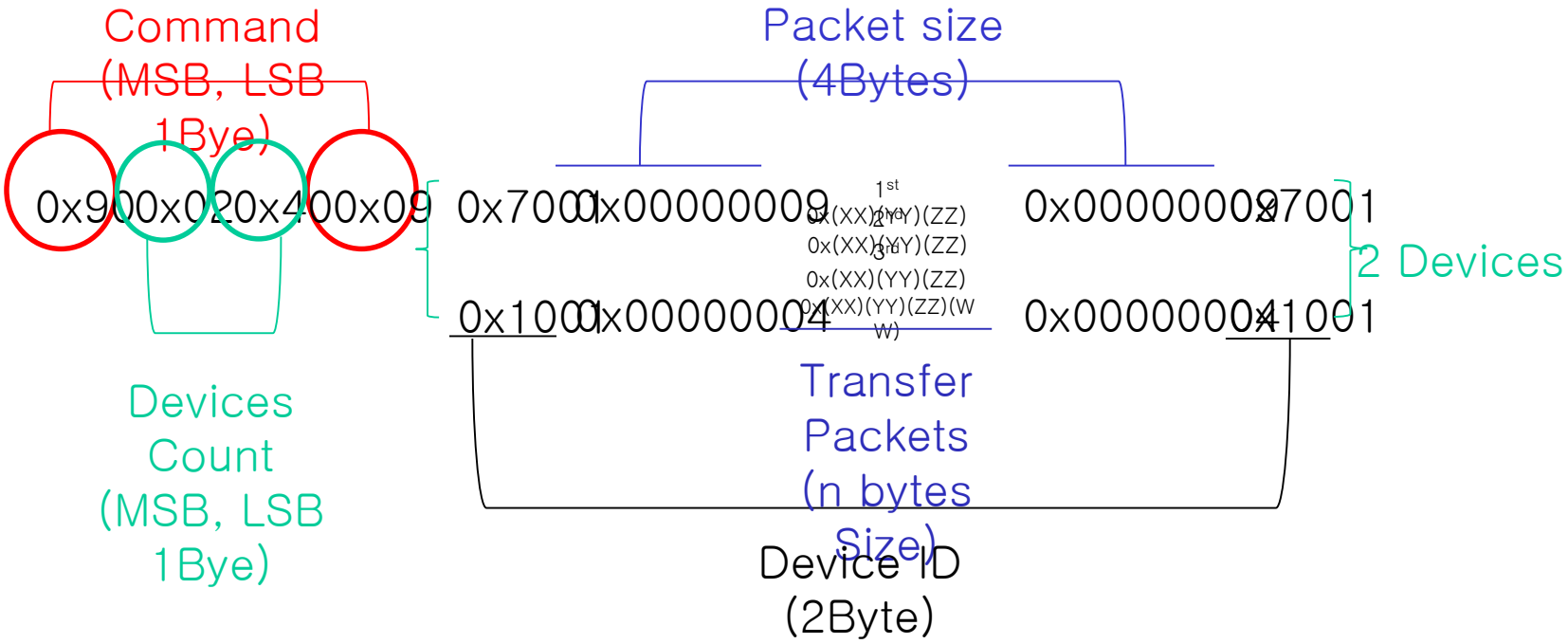
Haptics Issues of multimodal Interfaces in Virtual Training

- 9 DOF Sensor (Accelerometer/Gyroscope/Magnetometer)
- Vibration
- Reel Control Motor
- Haptic Actuator
- Pressure Sensor

Multimodal Interfaces Input Data Transfer for Haptic

- Input Type 2 Devices Data Transfer.

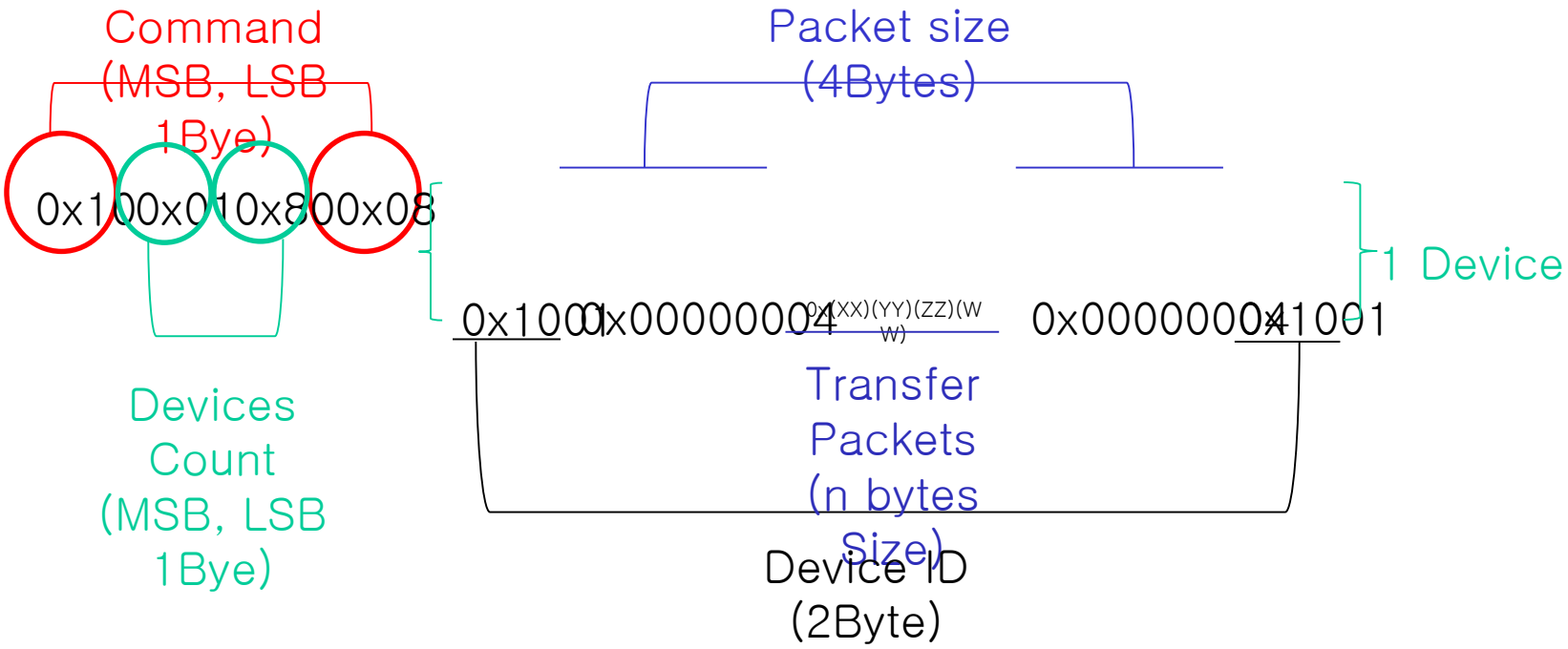
Touch(Device ID : 0x1001 – Haptic Gluves) , Sight / Hearing / Touch (Device ID : 0x7001 – 9DOF Sensor[Headtracking Sensor])



Multimodal Interfaces Output Control Transfer for Haptic

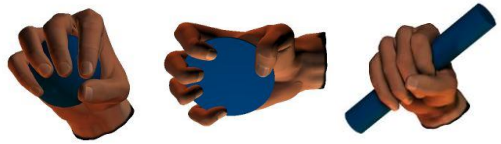
- Output Type 1 Device Control Transfer.

Touch(Device ID : 0x1001 – Haptic Gluves)



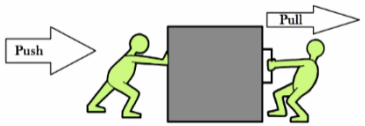
Haptics Issues of multimodal Interfaces in Virtual Training

Touch / Grasp



Volume

Weight



Strength of pow

Synchronization Issues of multimodal interfaces

Synchronization Issues of multimodal Interfaces in Virtual Training

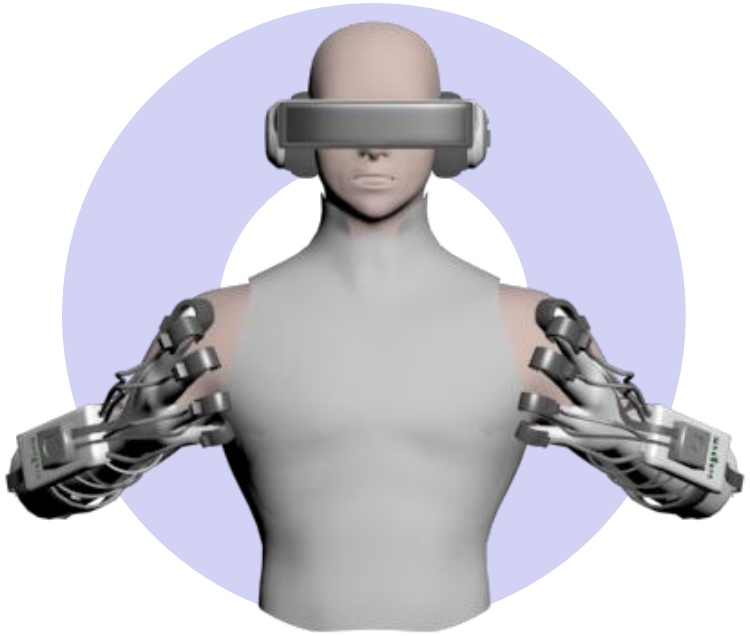
Input

- Voice Recognition
- Gesture Recognition
- Headtracking
- Motion Tracking
- Spatial Sensing

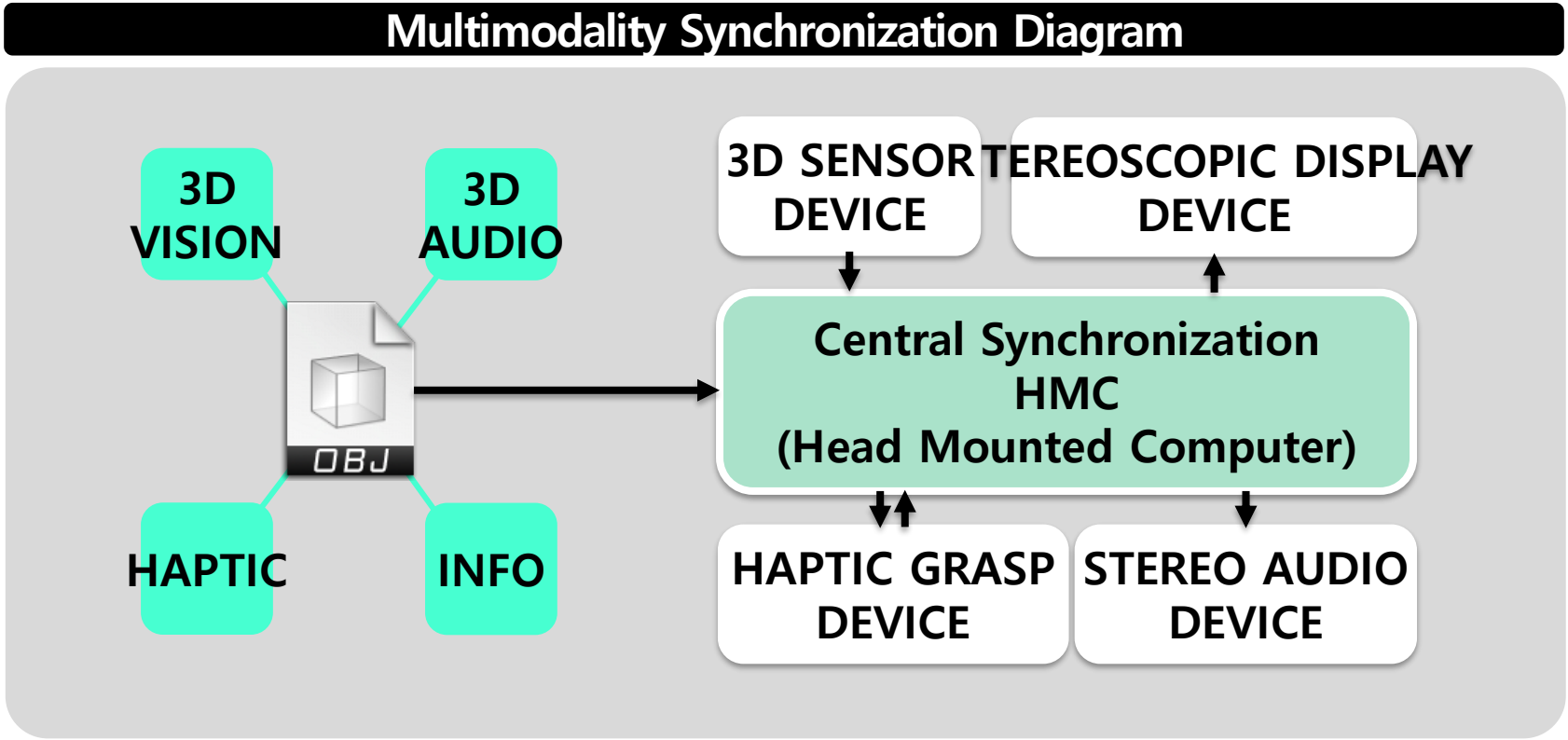
Output

- VR / AR
- MR / XR
- Binaural Effect 3D Audio
- Touch, Volume, Weight

synchronization

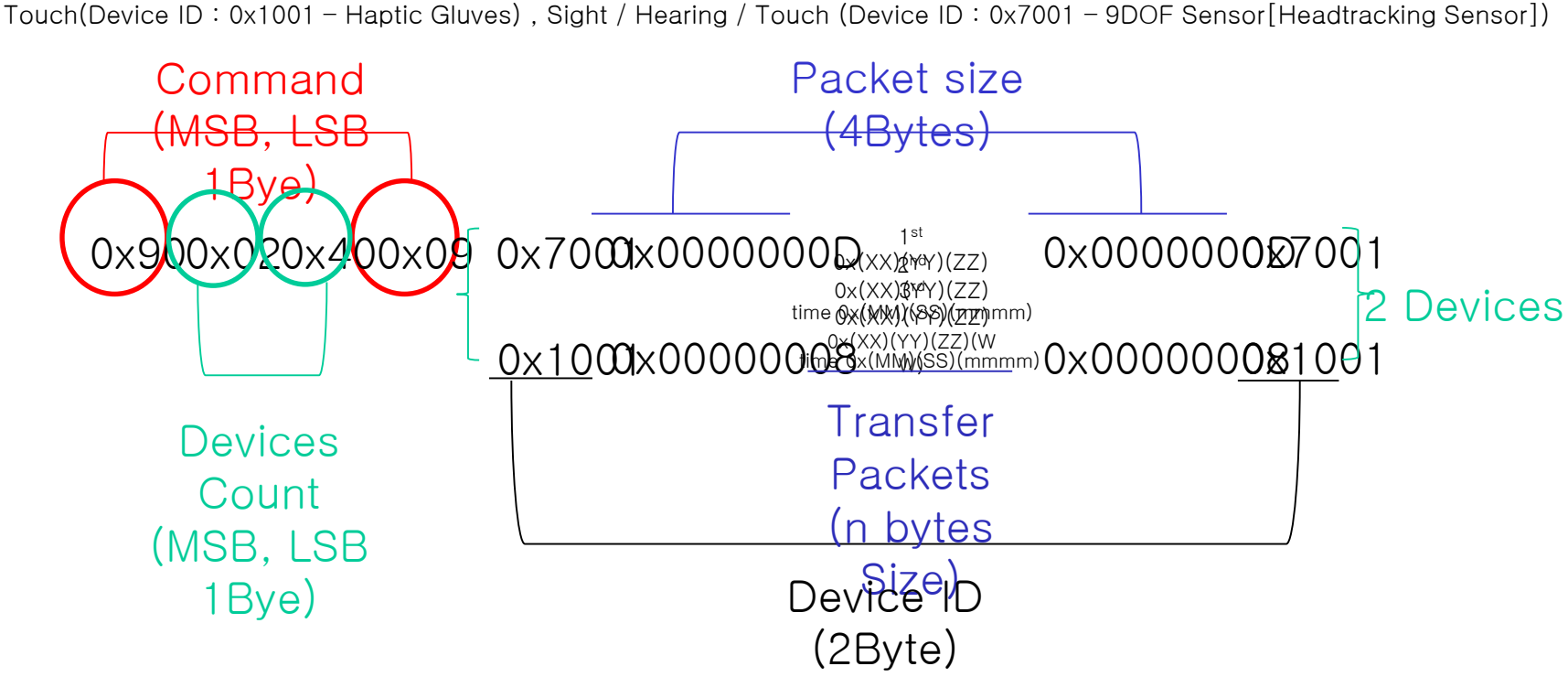


Synchronization Issues of multimodal Interfaces in Virtual Training



Multimodal Interfaces Synchronization Example - Input

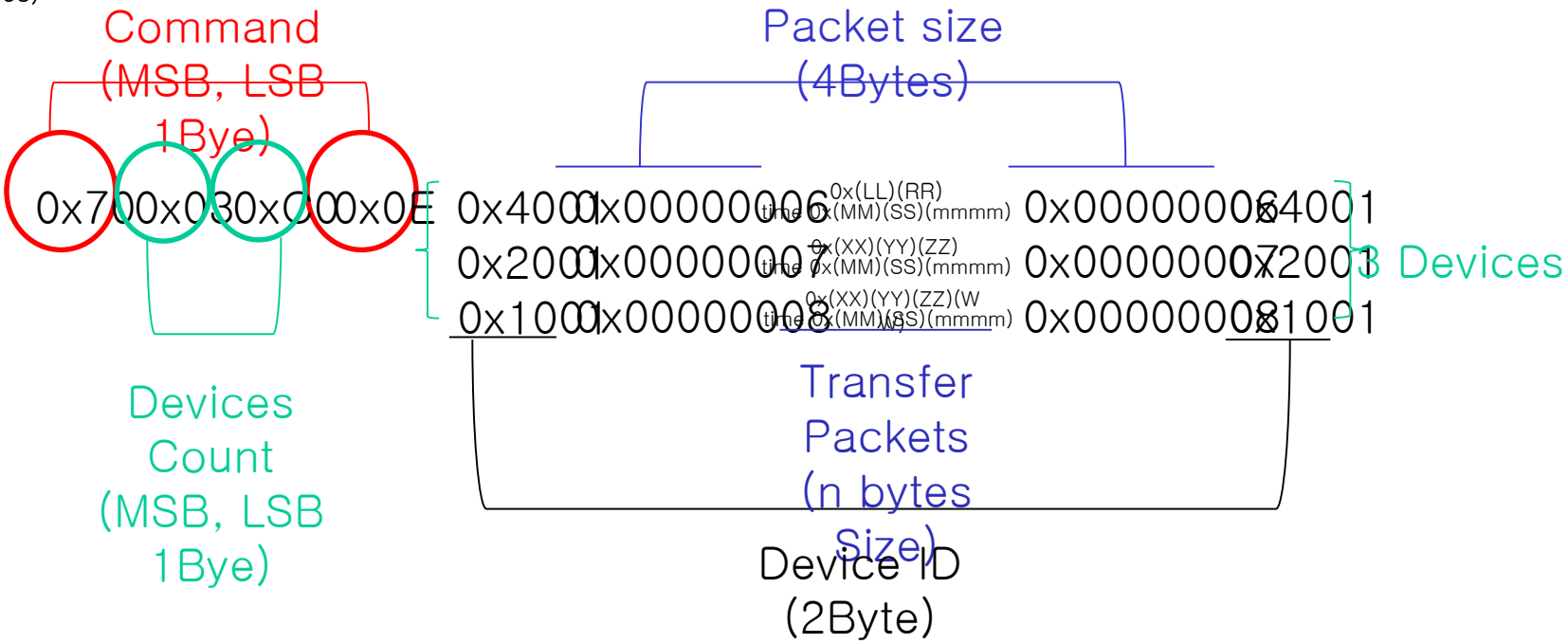
- Input Type 2 Devices Data Transfer.



Multimodal Interfaces Synchronization Example - Output

- Output Type 3 Devices Data Transfer Sync.

Sight(Device ID : 0x4001 – 3D Display) / Hearing(Device ID : 0x2001 – 3D Audio Headphone) / Touch(Device ID : 0x1001 – Haptic Gluves)

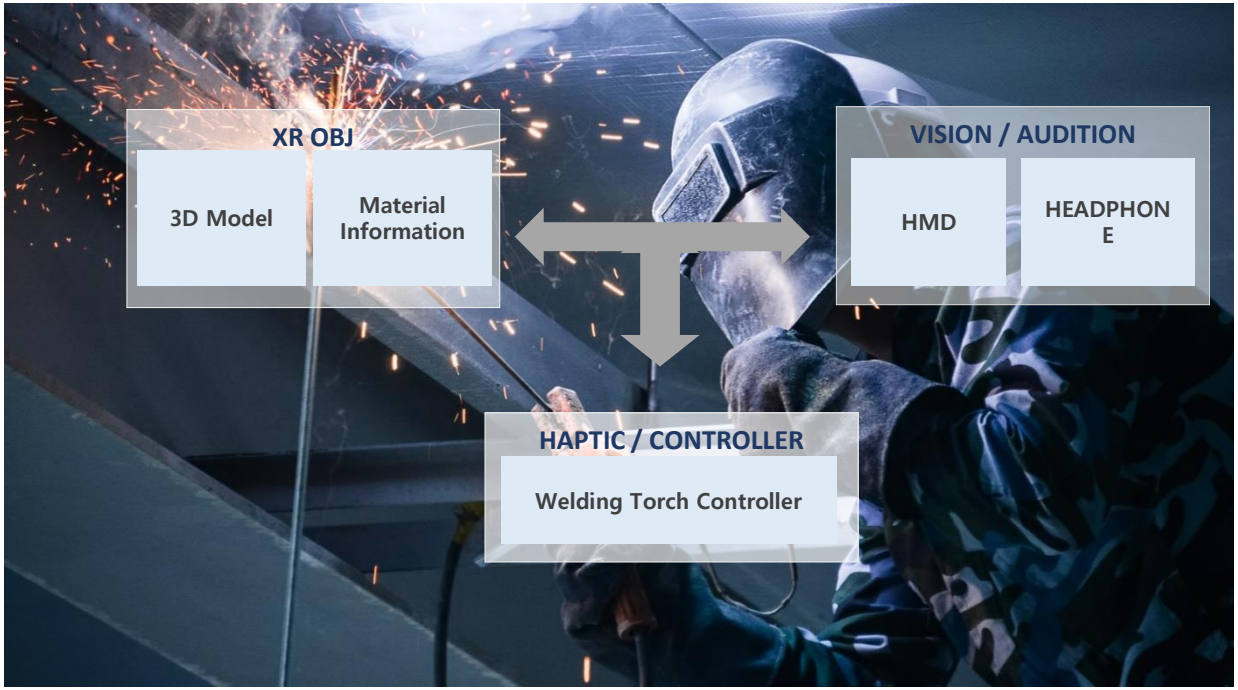


Synchronization Issues of multimodal Interfaces in Virtual Training

Use cases

- Standalone training
  - Welding
  - Driving
    - Car
    - Flight
    - Ship / Train
    - Heavy equipment
  - Space / Submersion
  - Manufacturing
  - Weapon
  - Surgery
- Multi-training
  - Sports
  - War
  - 1:N / N:N Training

Use case : Welding



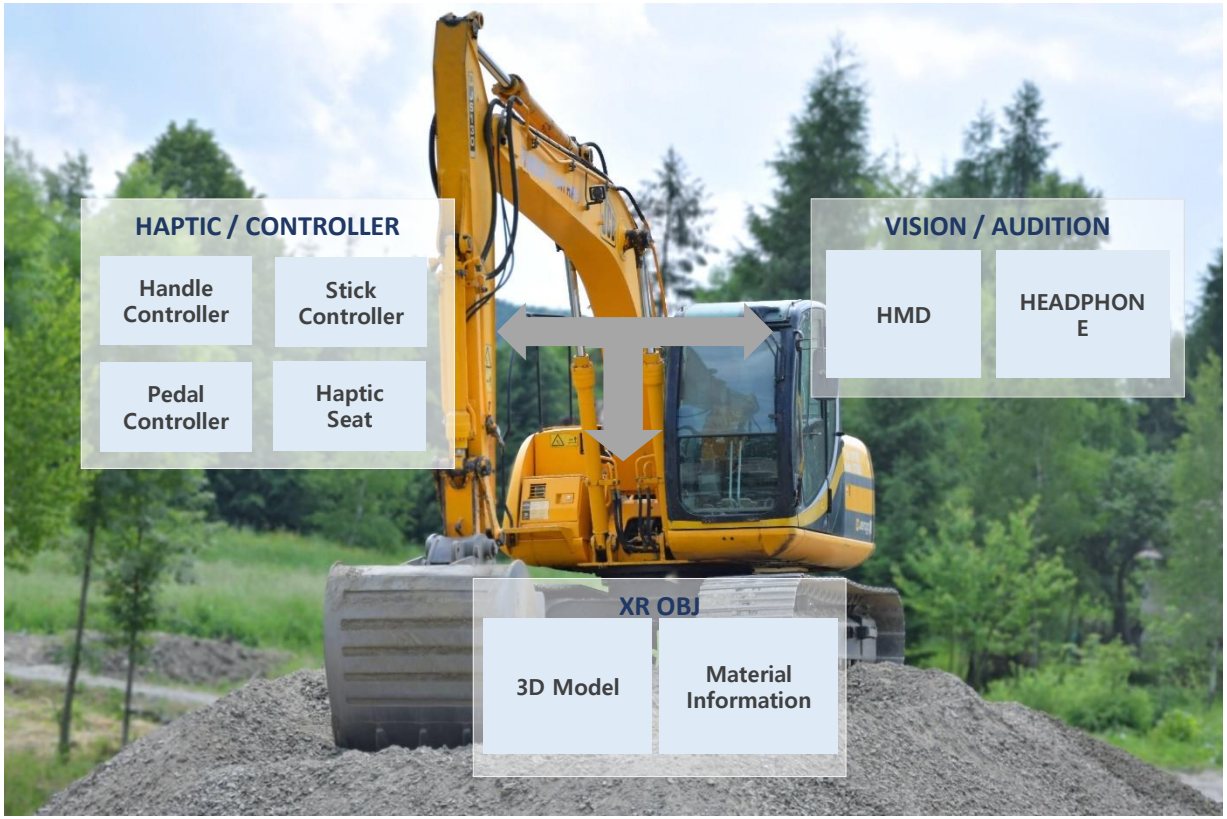
Welding

- Device (Vision, Audition, Haptic, Controller)
  - Capability
  - Connection
  - Session

- Vision (Rendering)
- Audition
- Haptics
- Controller

- Async or Synchronization of the three interfaces
  - ✓ Time Marking Async / Sync.
  - ✓ Event Marking Async / Sync.

Use case : Excavators



Excavators

- Device (Vision, Audition, Haptic, Controller)
  - Capability
  - Connection
  - Session

- Vision (Rendering)
- Audition
- Haptics * 1
- Controller*3

- Async or Synchronization of the three interfaces
  - ✓Time Marking Async / Sync.
  - ✓Event Marking Async / Sync.

Device Vision

- Capability
  - Stereoscopic Duel Display
- Connection
  - 9DOF MEMS sensor
- Session

Device Audition

- Capability
  - 3D Stereo speakers.
- Connection
  - 9DOF MEMS Sensor
- Session
  - HRTF Binaural Effect

Device Haptic

- Capability
  - 3D Pressure / Capacitive output
- Connection
  - Controller Device
- Session
  - Human Muscle Info
  - Human Skin Info

Device Controller

- Capability
  - 3D Gesture Sensing
- Connection
  - Haptic Device
- Session
  - 9DOF MEMS Sensor

Node Definition

- Nodes of Device (Vision, Audit, haptic)
  - Capability
  - Connection
  - Session

- Node of Vision (Rendering)
- Node of Audition
- Node of Haptics

- Node of Async. or Synchronization of the three interfaces
  - ✓ Time Marking Async / Sync.
  - ✓ Event Marking Async / Sync.

Vision Nodes

- Node of Vision
  - Id
  - Description
  - Raw signals (encoding)
  - Start time – end time
  - Position
  - Direction + aperture
  - Collision

<VTS-VISION-Cmd id= "init" description= "Initialize Display device" sync="true" output="true" deviceCount='1' requestDevice=" {deviceID:1, deviceName: 'device1', x='0', y='0', z='0'}">
</ VTS-VISION-Cmd >

< VTS-VISION-Data id= "device1" description= "Transfer Data" sync="false" timeStamp="2020/05/20:12:00:00:000" timeLimit="2020/05/20:12:00:00:010" deviceName= "device1" requestSession= "" sessionOptions= "{}">
[data]
</ VTS-VISION-Data >

Audition Nodes

- Node of Audition
  - Id
  - Description
  - Raw signals (encoding)
  - Start time – end time
  - Position
  - Direction + aperture
  - Collision

<VTS-AUDITION-Cmd id= "init" description= "Initialize Audition device" sync="true" output="true" deviceCount='1' requestDevice=" {deviceID:1, deviceName: 'device1', x='0', y='0', z='0'}">
</ VTS-AUDITION-Cmd >

< VTS-AUDITION-Data id= "device1" description= "Transfer Data" sync="false" timeStamp="2020/05/20:12:00:00:000" timeLimit="2020/05/20:12:00:00:010" deviceName= "device1" requestSession= "" sessionOptions= "{}">
[data]
</ VTS-AUDITION-Data >

Haptics Nodes

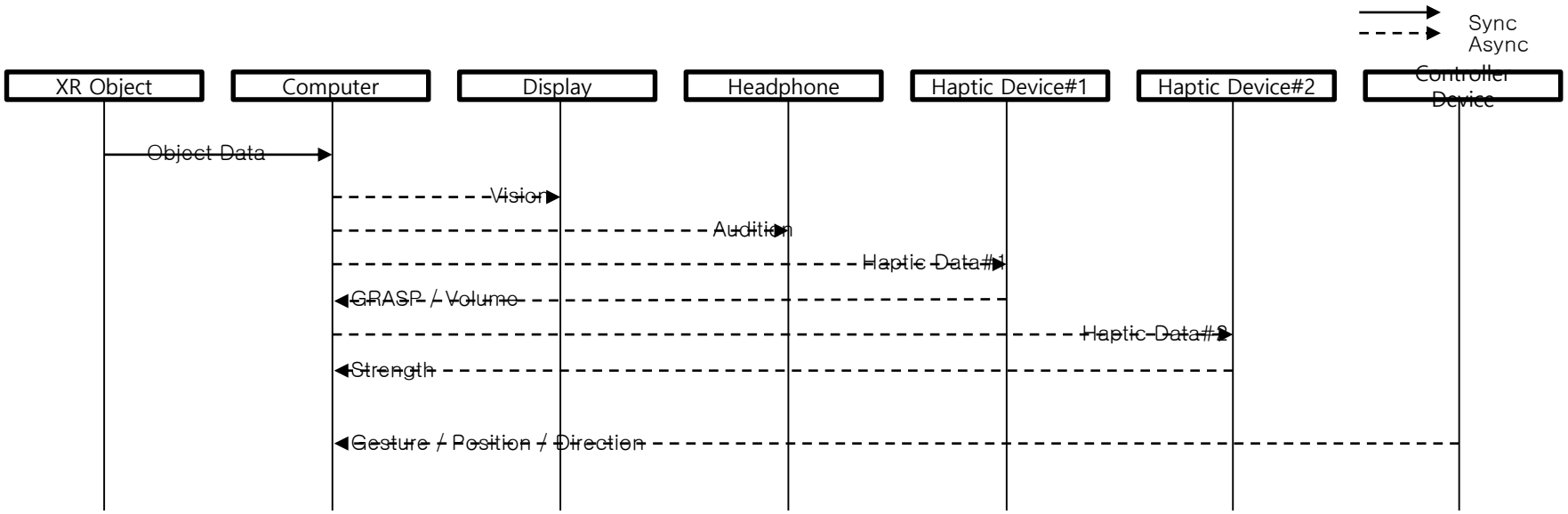
- Node of Haptics
  - Id
  - Description
  - Raw signals (encoding)
  - Start time – end time
  - Position
  - Direction + aperture
  - Collision

<VTS-HAPTIC-Cmd id= "init" description= "Initialize Haptic device" sync="true" output="true" deviceCount='1' requestDevice=" {deviceID:1, deviceName: 'device1', strength='0', map='0'}">
</ VTS-HAPTIC-Cmd >

< VTS-HAPTIC-Data id= "device1" description= "Transfer Data" sync="false" timeStamp="2020/05/20:12:00:00:000" timeLimit="2020/05/20:12:00:00:010" deviceName= "device1" requestSession= "" sessionOptions= "{}">
[data]
</ VTS-HAPTIC-Data >

Synchronization Nodes

- Node of Synch.



New Work Items

- Virtual training contents
- Vision (Rendering) interface
- Audition interface
- Haptics/Controller interface
- Synchronization of the three interfaces
  - ✓ Time
  - ✓ Event

Scope of multimodal interface

- Concepts of multimodal interface in virtual training system
- Requirements for 3D simulation of virtual training system
- How to represent multimodal interfaces in a virtual training system
- How to organize multimodal interfaces
- How to define an abstract model for representing multimodal interfaces in a virtual training system
- How to define a system architecture for representing multimodal interfaces in a virtual training system
- How to use multimodal interfaces in a virtual training system
- How to synchronize multimodal interfaces in a virtual training system

What to focus on

- How to represent multimodal interface in a virtual training system
  - **Visual interface**
  - **Control interface**
  - **Audio interface**
  - **Synchronization of three interfaces**
- What to represent about multimodal interfaces
  - Visual properties, Control properties, Audio properties,
- What to do
  - Control (device) interfaces in VTS
  - Virtual training simulation using multimodal interface
- Reason for doing
  - Providing natural interaction with VTS
  - Improve training effect

Contents of specification

1. Scope
2. Normative references
3. Terms, definitions, acronyms, and abbreviations
4. Concepts
   - Overview
   - Virtual training systems
   - Components of virtual training systems
   - Interaction interface for virtual training systems
5. Multimodal interface in VTS
   - Overview
   - Visual Interface
   - Control Interface
   - Audio Interface
   - Synchronization
6. Node Definition
   - Overview
   - Nodes
7. Conformance

Q&A