ISO/IEC JTC 1/SC 24/WG 9 "Augmented reality continuum concepts and reference model"
Convenorship: **KATS**
Convenor: **Kim Gerard Jounghyun Mr**

Synchronization of Multimodal Data in MAR

Document type	Related content	Document date	Expected action
Meeting / Presentation	Meeting: VIRTUAL 21 Jul 2021	2021-11-03	

Synchronization of Multimodal Data in MAR

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

13 July – 5 August 2021

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

Concepts

Properties of multimodal data in MAR

- 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 data in MAR

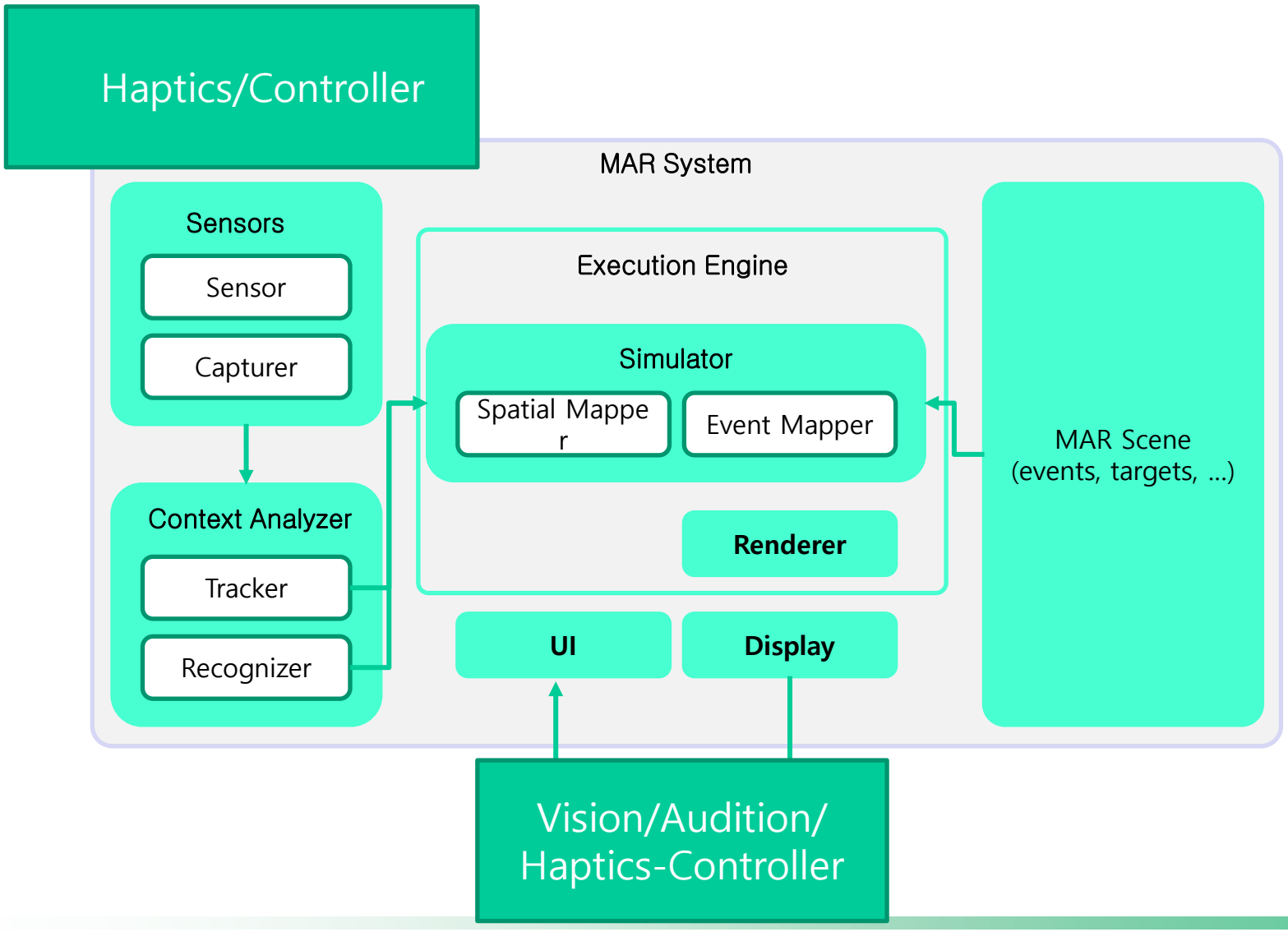
- Vision (Rendering): A large number of perceptual primitives.
  - ✓ Cortical visual pathways: where(dorsal) stream, what(ventral) stream
  - ✓ What: the analysis of form and color
  - ✓ Where: the analysis of motion and spatial relations
  - ✓ Perception: color, space, motion
  - ✓ HMD(head mounted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

Properties of multimodal data in MAR

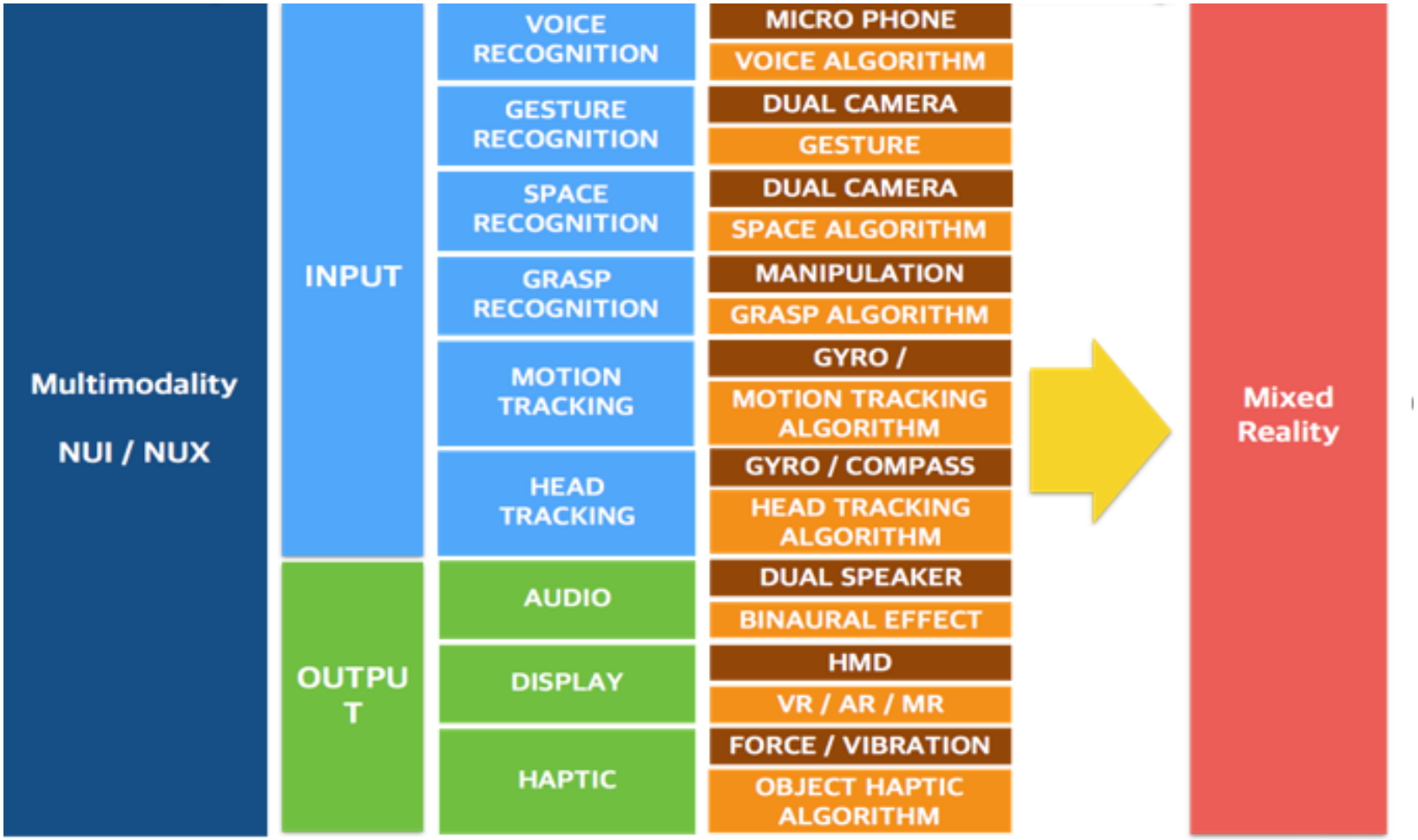
- Audition: the localization of sound in 3D spaces.
  - ✓ auditory analysis of scenes
  - ✓ Music and speech perception
  - ✓ Sound feedback

Properties of multimodal data in MAR

- 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


Haptics/Controller
MAR System
Sensors
Sensor
Capturer
Context Analyzer
Tracker
Recognizer
Execution Engine
Simulator
Spatial Mappe r
Event Mapper
Renderer
UI
Display
MAR Scene (events, targets, ...)
Vision/Audition/ Haptics-Controller

Multimodal data in MAR



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

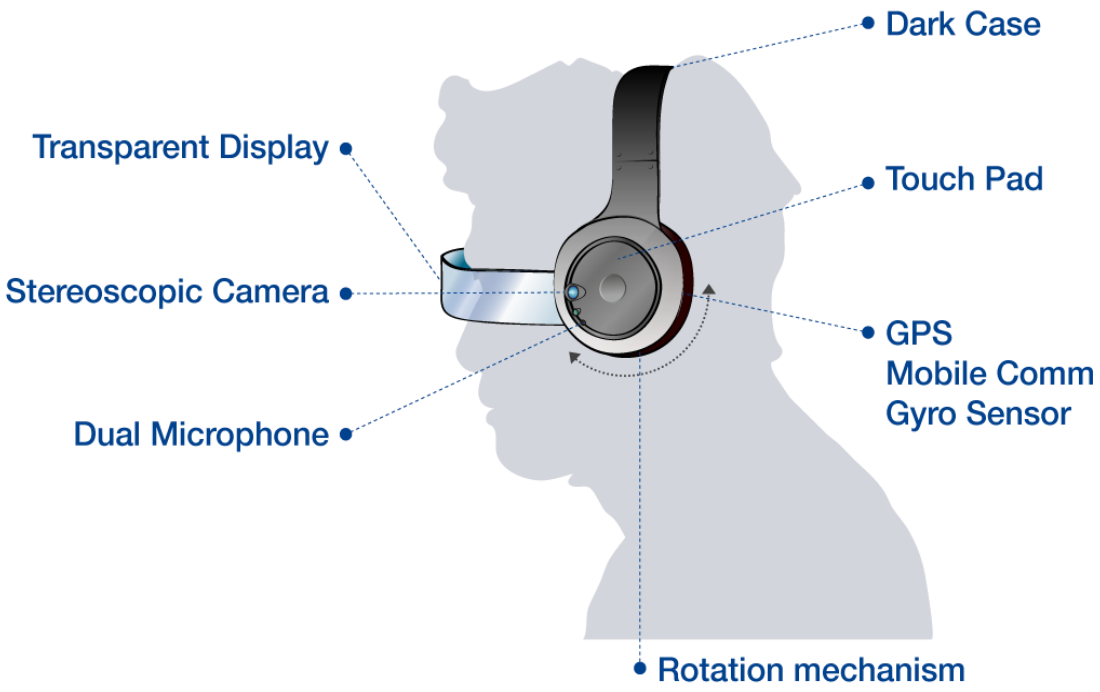
Multimodal Data Format Structure

- Multimodality Interfaces Protocol Input/Output Format

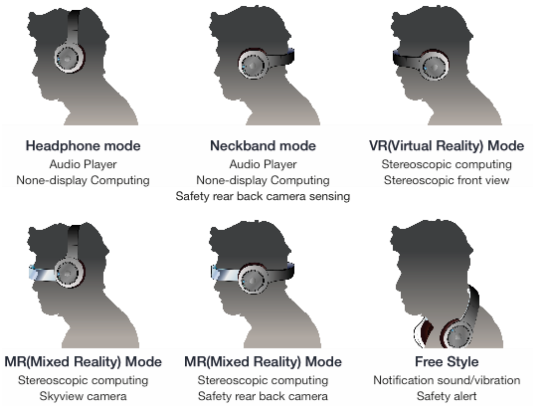
I/O	Command (MSB First 1Bye) xSHTSTRx								Devices Count (MSB First 1Bye)	Devices Count (LSB First 1Bye)	Command (LSB 1Bye)							0~n Device ID (2Bytes)	0~n Packet size (4Bytes)	Packets (Packet Size)	0~n Packet size (4Bytes)	0~n Device Number (2Bytes)
	I/O	Sight	Hearing	Touch	Smell	Taste	Reserve	CmdFlag	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

Vision Interface Issues

Vision Issues of multimodal Interfaces in Virtual Training

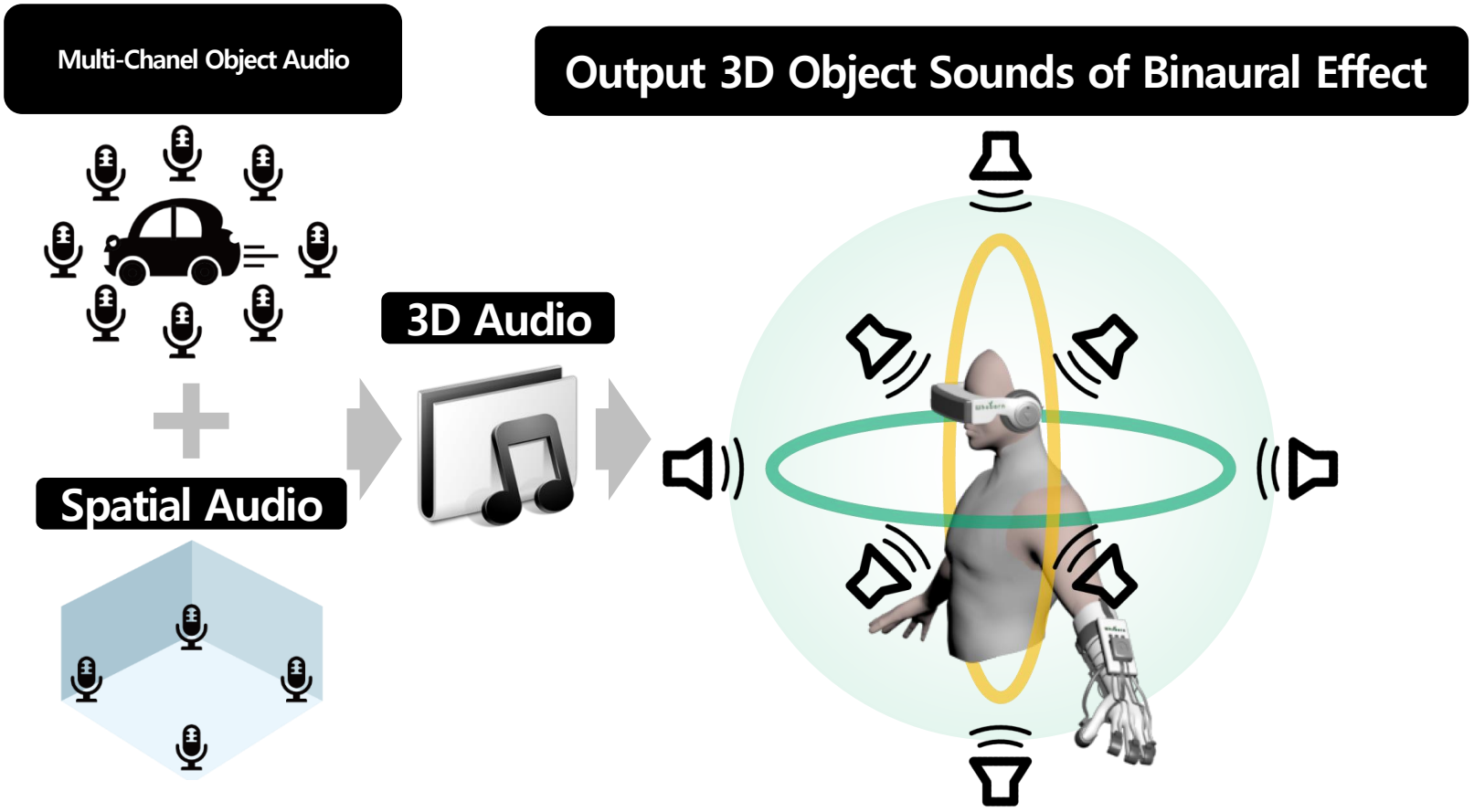


Vision Issues of multimodal Interfaces in Virtual Training



Audition Interface Issues

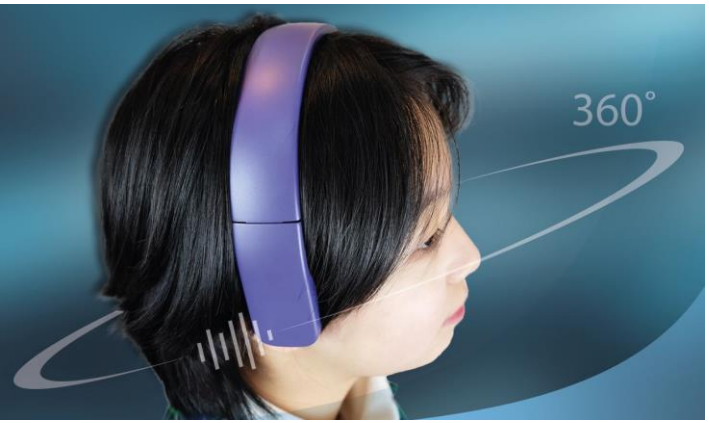
Audition Issues of multimodal Interfaces in Virtual Training



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

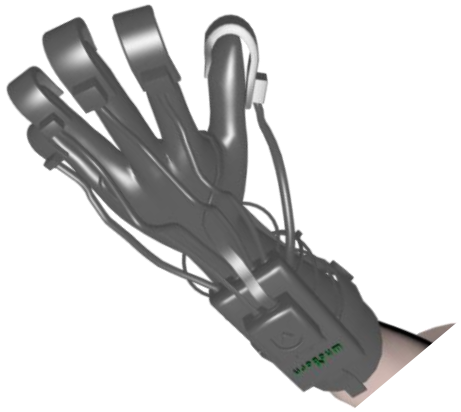
Audition Issues of multimodal Interfaces in Virtual Training



Haptics/Controllers Interface Issues

Haptics Issues of multimodal Interfaces in MAR

- 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 MAR

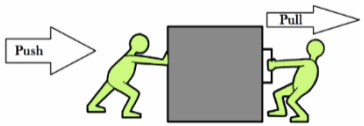
- 9 DOF Sensor (Accelerometer/Gyroscope/Magnetometer)
- Vibration
- Reel Control Motor
- Haptic Actuator
- Pressure Sensor

Haptics Issues of multimodal Interfaces in MAR

Touch / Grasp

Volume

Weight



Strength of pow

Synchronization Issues of multimodal interfaces

Synchronization Issues of multimodal Data in AMR

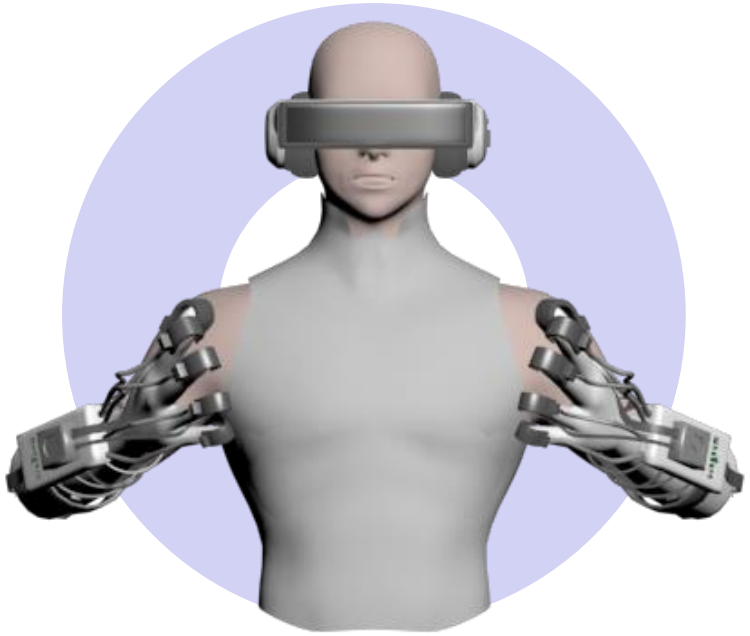
Input

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

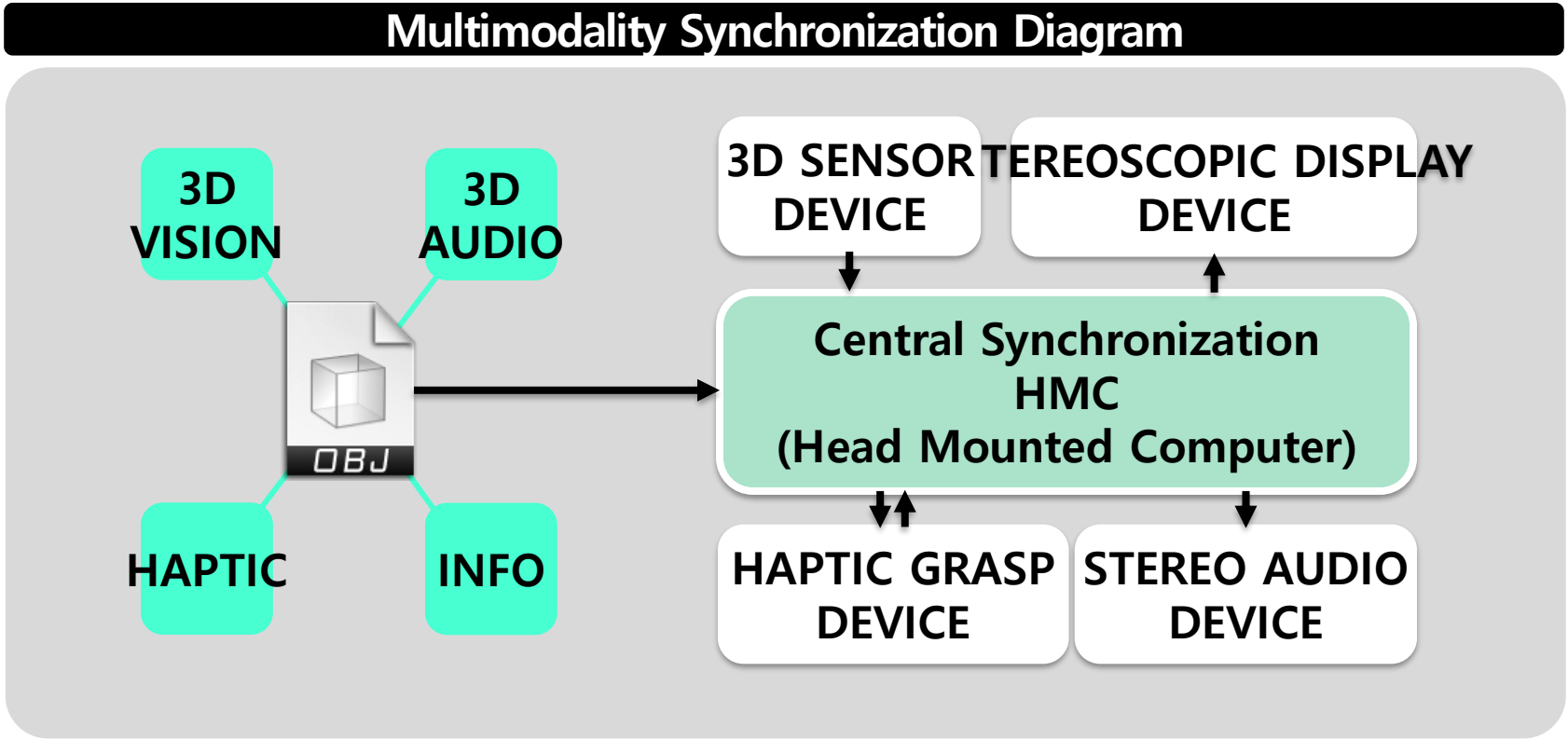
Output

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

synchronization



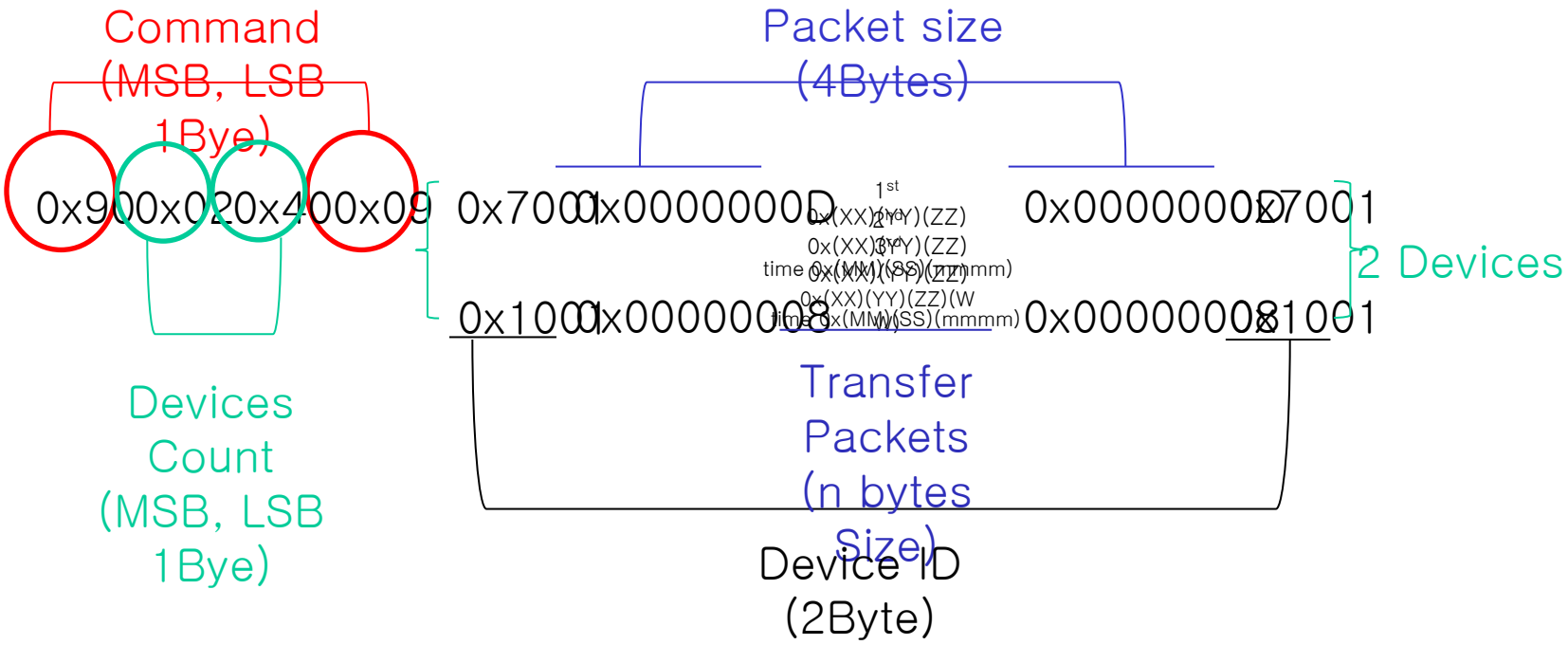
Synchronization Issues of multimodal Data in MAR



Multimodal Data Synchronization Example - Input

- Input Type 2 Devices Data Transfer.

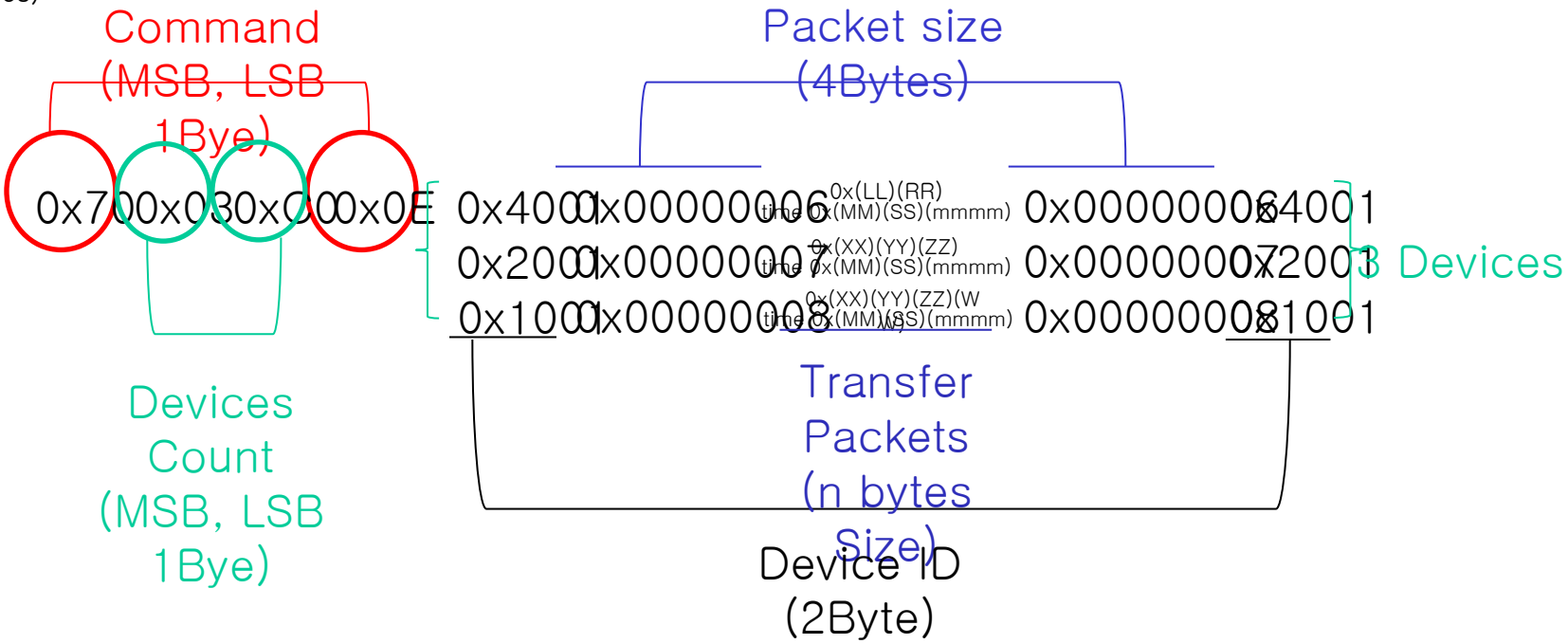
Touch(Device ID : 0x1001 – Haptic Gluves) , Sight / Hearing / Touch (Device ID : 0x7001 – 9DOF Sensor[Headtracking Sensor])



Multimodal Data 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)

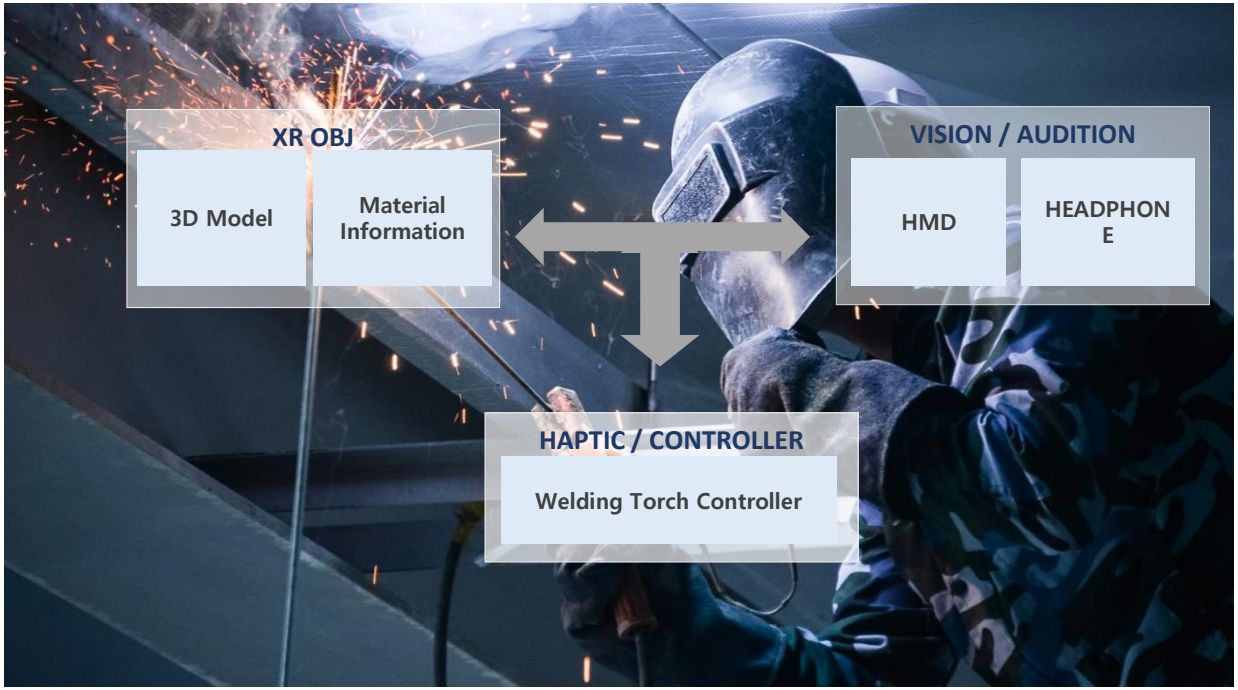


Synchronization Issues of multimodal Data in MAR

Use cases

- Standalone MAR system
  - Welding
  - Driving
    - Car
    - Flight
    - Ship / Train
    - Heavy equipment
  - Space / Submersion
  - Manufacturing
  - Weapon
  - Surgery

Use case : Welding



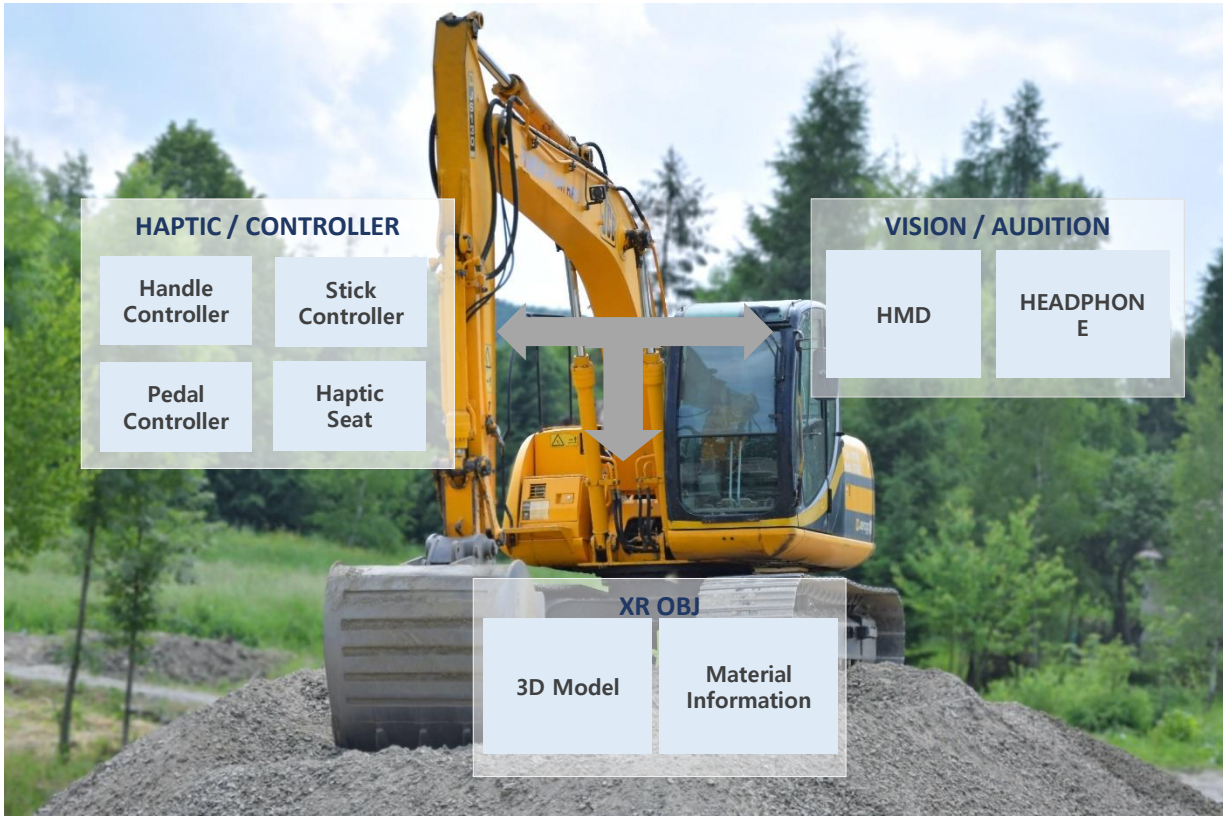
Welding

- 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

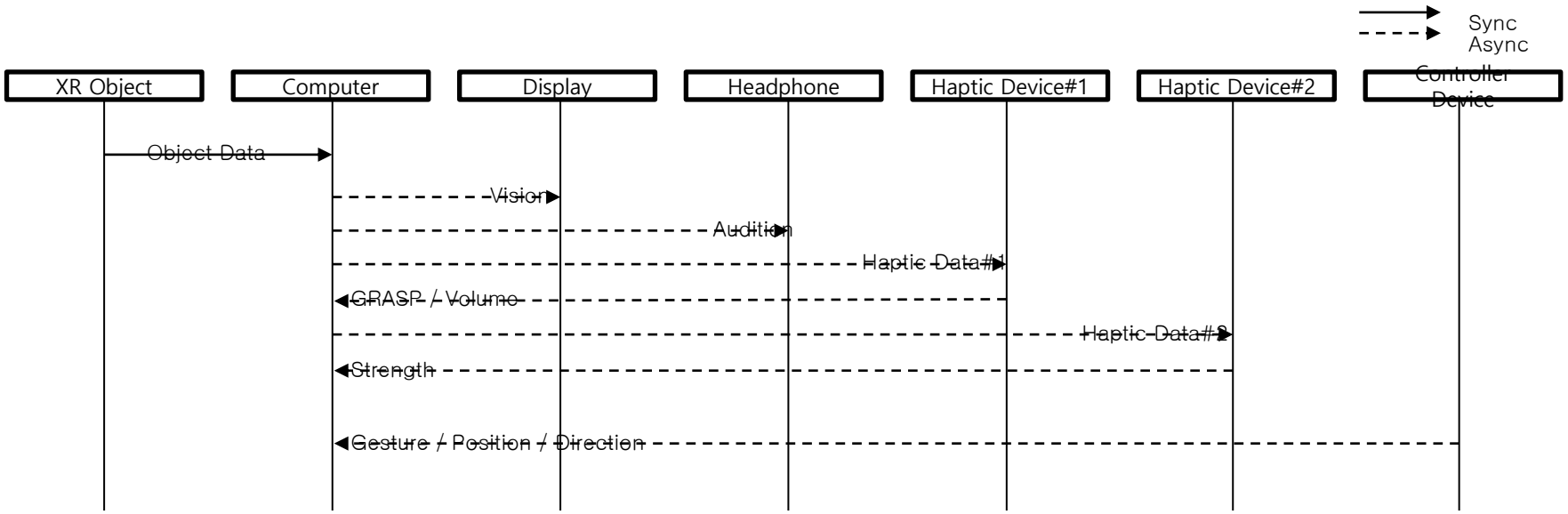
- Node of Haptics
  - Id
  - Description
  - Raw signals (encoding)
  - Start time – end time
  - Position
  - Direction + aperture
  - Collision

<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 MAR 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 MAR
  - Simulation using multimodal interface
- Reason for doing
  - Providing natural interaction with a MAR system
  - Improve training effect

Contents of specification

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

Q&A