Recommendation

ITU-T H.430.6 (09/2023)

SERIES H: Audiovisual and multimedia systems

Infrastructure of audiovisual services – Telepresence, immersive environments, virtual and extended reality

Media transport protocols, signalling information of haptic transmission for immersive live experience (ILE) systems

ITU-T H-SERIES RECOMMENDATIONS

Audiovisual and multimedia systems

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T H.430.6

Media transport protocols, signalling information of haptic transmission for immersive live experience (ILE) systems

Summary

Immersive live experience (ILE) systems may handle haptic information such as vibrotactile and kinaesthetic actions, for increasing more immersiveness in addition to audio and video. Haptic information should be transmitted synchronously with audio, video, and lighting information. Recommendation ITU-T H.430.6 identifies media transport protocol and signalling information of haptic transmission for immersive live experience (ILE) systems, in order to transmit haptic information synchronously to provide ILE services.

Keywords

Haptic, ILE, kinaesthetic actions, media transport protocol, signalling information, vibrotactile.

---

* To access the Recommendation, type the URL https://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Recommendation ITU-T H.430.6

Media transport protocols, signalling information of haptic transmission for immersive live experience (ILE) systems

1 Scope

Immersive live experience (ILE) systems may handle haptic information such as vibrotactile and kinaesthetic actions, for increasing more immersiveness in addition to audio and video. Haptic information should be transmitted synchronously with audio, video, and lighting information, however [ITU-T H.430.4] does not clearly specify the signalling information about haptic transmission.

This Recommendation identifies media transport protocol and signalling information for haptic transmission over immersive live experience (ILE) systems.

The scope of this Recommendation includes:

– Requirements of haptic transmission for ILE systems;

– Media transport protocol for haptic transmission;

– Signalling information for haptic transmission.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T H.264]   Recommendation ITU-T H.264 (V14) (2021), *Advanced video coding for generic audiovisual services*.

[ITU-T H.265]   Recommendation ITU-T H.265 (V9) (2023), *High efficiency video coding*.

[ITU-T H.266]   Recommendation ITU-T H.266 (V3) (2023), *Versatile video coding*.

[ITU-T H.430.4]   Recommendation ITU-T H.430.4 (2019), *Service configuration, media transport protocols, signalling information of MPEG media transport for immersive live experience (ILE) systems*.

[ISO/IEC 14496-12] ISO/IEC 14496-12:2022, *Information technology – Coding of audio-visual objects – Part 12: ISO base media file format*.

[ISO/IEC 23005-2] ISO/IEC 23005-2:2018, *Information technology – Media context and control – Part 2: Control information*.

[ISO/IEC 23005-3] ISO/IEC 23005-3:2019, *Information technology – Media context and control – Part 3: Sensory information*.

[ISO/IEC 23008-1] ISO/IEC 23008-1:2023, *Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 1: MPEG media transport (MMT)*.

[ISO/IEC 23009-1]	ISO/IEC 23009-1:2022, *Information technology – Dynamic adaptive streaming over HTTP (DASH) – Part 1: Media presentation description and segment formats.*
[ISO 9241-910]	ISO 9241-910:2011, *Ergonomics of human-system interaction – Part 910: Framework for tactile and haptic interaction.*
[ISO 9241-920]	ISO 9241-920:2009, *Ergonomics of human-system interaction – Part 920: Guidance on tactile and haptic interactions.*

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 haptics [ISO 9241-910]: Haptics consists of touch (tactile/cutaneous) and kinaesthesis (kinaesthetic), which are sensory and/or motor activity based in the skin, muscles, joints and tendons.

3.1.2 immersive live experience (ILE) [b-ITU-T H.430.1]: A shared viewing experience which stimulates emotions within audiences at both the event site and remote sites, as if the ones at remote sites wandered into substantial event site and watched actual events in front of them, from high-realistic sensations brought by a combination of multimedia technologies such as sensorial information acquisition, media processing, media transport, media synchronization, and media presentation.

3.1.3 kinaesthetic [ISO 9241-910]: Kinaesthetic appertains to kinaesthesis.

3.1.4 kinaesthesis [ISO 9241-910]: Kinaesthesis is sense and motor activity based on muscles, joints and tendons.

3.1.5 tactile [ISO 9241-910]: Tactile appertains to touch.

3.1.6 vibrotactile [ISO 9241-910]: Vibrotactile is vibration-based stimulation of the skin.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

3D	Three Dimensional
AAC	Advanced Audio Coding
DASH	Dynamic Adaptive Streaming over HTTP
ILE	Immersive Live Experience
MFU	Media Fragmentation Unit
MMT	MPEG Media Transport
MMTP	MMT Protocol
MPD	Media Presentation Description
MPEG	Moving Picture Experts Group
MPI	Media Presentation Information
PI	Presentation Information

TS   Transport Stream

UTC   Coordinated Universal Time

XML   Extensible Mark-up Language

5   Conventions

In this Recommendation:

–   The first line of XML is written for document reputation, the keywords "version" is a required attribute, and the keywords "encoding" indicate a coding scheme which tells the parsing engine what character set the current document is using.

–   The keywords "is required to" indicate a requirement which must be strictly followed and from which no deviation is permitted, if conformance to this Recommendation is to be claimed.

6   Requirements to haptic transmission for ILE systems

Haptic information, including vibrotactile, may increase immersiveness and reality in immersive live experience (ILE) services. At present there are several standards such as [ITU-T H.264], [ITU-T H.265], [ITU-T H.266] and Moving Picture Experts Group (MPEG)-2 advanced audio coding (AAC) for transmitting audio and video in ITU-T and ISO/IEC JTC1 (MPEG). However, no standards for haptic information transmission were yet considered. ILE services can provide high-realistic sensations at the viewing sites by simultaneous transmission of audio, video lighting and other information. In addition, simultaneous transmission of haptic information with other information increases more realistic sensation at the viewing sites.

In order to transmit haptic information from the event site to the viewing sites, media transport protocols and signalling information of haptic transmission for ILE systems needs to be clarified.

6.1   Haptic information

Haptic information normally includes the following:

1)   Information on haptic sensors

–   Type of sensor (tactile, vibrotactile, kinaesthetic)

–   ID of the object the sensor is attached to

–   Three dimensional (3D) position sensed, in general sensor position attached to the object

–   3D direction sensed, in general sensor direction attached to the object

2)   Tactile sensor information

–   Machinery information

–   Temperature information

3)   Vibrotactile sensor information

–   Frequency information

–   Intensity information

4)   Kinaesthetic sensor information

–   Force information

–   Torque information

–   Stiffness information

The above information is the basis for describing haptics, and more information might be needed to be transmitted for reconstructing the haptic atmosphere at the viewing sites.

Tactile sensor information, vibrotactile sensor information and kinaesthetic sensor information are defined in [ISO/IEC 23005-2], [ISO/IEC 23005-3], [ISO 9241-910], and [ISO 9241-920] which include the extensible mark-up language (XML) schema and syntax. In addition, HapticSampleEntry in the ISOBMFF standard [ISO/IEC 14496-12] and encoded haptics to the media presentation description (MPD) segments in an upcoming release of the MPEG-dynamic adaptive streaming over HTTP (DASH) streaming standard [ISO/IEC 23009-1] are under study in ISO/IEC JTC SC29.

This Recommendation can refer to these existing specifications for representing haptic information. However, 3D position data for the source of haptic information are not defined in these specifications, so this Recommendation needs to define special information about haptic sensors.

6.2   Spatial information

Haptic information is gathered by sensors in event sites, and the sensors are allocated several points that have spatial information such as the 3D position. In order to reconstruct the atmosphere of the event site at the viewing sites, the spatial information of sensors at the event site needs to be transferred. The information can be transferred by using reference space, as defined in [ITU-T H.430.4]. The reference space is shown in Figure 1.
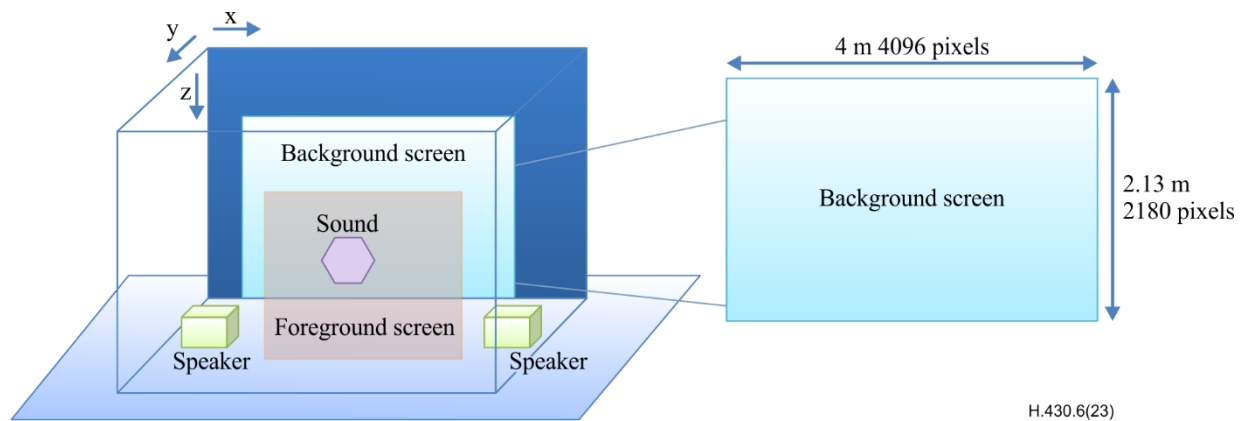
Figure 1 – Example of reference space describing the location and size of equipment

7   Media transport protocols for haptic transmission

Haptic information might be transmitted over several media transport protocols, such as multimedia data in the MPEG media transport (MMT) protocol (MMTP) and MPEG-2 transport stream (TS). Vibration information might be considered as the low frequency audio signals, thus this Recommendation aims to use existing media transport for audio such as MMT, MPEG-2 TS, and other multiplex formats for transmitting vibrotactile sensor information.

ILE systems using haptic information are required to synchronize haptic signals with audio and video signals for increasing immersive sensation for users. As one of the media transport protocols, MMTP can carry multimedia data synchronously. [ISO/IEC 23008-1] specifies the encapsulation format, but there are some variations of the media fragmentation unit (MFU) format. Since haptic information can be transmitted very similarly with audio signals, the encapsulation method defined in [ITU-T H.430.4] is utilized. When MPEG-2 TS is used, haptic information can be encapsulated and transmitted by one of the audio channels.

8   Signalling information for haptic transmission

In order to handle assets including spatial information for reconstructing haptic information at viewing sites, the following constraints are applied to the signalling information for media descriptors.

8.1 Environment descriptor

To reconstruct haptic information at viewing sites of ILE services, a presentation information (PI) content in media presentation information (MPI) tables is required to carry the spatial environment information. To specify the syntax of the environment descriptor of several sensors, which gather information described in clause 6.1, the environment descriptor defined in [ITU-T H.430.4] can be utilized.

Furthermore, other environmental information such as temperature, humidity, and pressure need to be gathered by tactile sensors in order to make the viewing sites as real as the event site.

In general, the temperature and pressure can be gathered by the sensors in the atmosphere or water. Temperature can be perceived by a temperature sensor in the unit of Celsius or Fahrenheit. Humidity represents the proportion of water in the air, which can be perceived by the humidity sensor in the unit of percentage, and the value of it indicates the degree of dryness and wetness. Pressure refers to the weight of the atmospheric column per unit area, which can be perceived by the atmospheric pressure sensor in the unit of hectopascal (hPa), and the value of it is closely related with weather and the altitude. In addition, if the event site is in the water, such as a swimming competition or underwater show, the temperature refers to water temperature, and the pressure stands for liquid pressure, its unit is N/mm^2 (Newton/millimetre square). The XML schema and syntax of the above environment descriptor are defined in [ISO/IEC 23005-3].

8.2 Haptic information and location descriptor

In order to reconstruct haptic information at viewing sites of ILE services, information of haptic sensors mentioned in clause 6.1, especially sensor location data, needs to be transmitted from the event site. Haptic information and location information should be transferred from the event site to viewing sites so that haptic information from the object can be accompanied by the displayed objects at the viewing sites. A haptic information and location descriptor is transferred on media transport protocols. In most cases, objects move around in the event site, so location information of the objects should be captured in each frame. In other words, haptic and location information are necessary to reproduce the event at the viewing sites.

As one of the media descriptors, the syntax of haptic information and location descriptor contains spatial information of sensors to reconstruct haptic phenomenon aligned with objects at the viewing sites. The information will be processed by media processing functions.

Sensors attached to objects are moving around in the event site, thus it is required to reconstruct the actual trajectory of the moving objects. For this purpose, the ILE systems need to transport tracking information of objects, as described in clause 7 of [ITU-T H.430.4]. The tracking information can support to specify the location of the haptic sensors.

The syntax of sensor_location_data and XML syntax sample are shown in Tables 1 and 2.

Table 1 – Syntax of sensor_location_data

Name	Data	Form	Repetition
sensor_location_data	Three-dimensional location information of each sensor		1
timecode	Time at which location info is consumed, which is denoted by coordinated universal time (UTC) for time synchronization between objects	String	1
object	Name of the object that anchors the sensor	String	0..*
sensor	Name of the sensor	String	1

Table 1 – Syntax of sensor_location_data

Name	Data	Form	Repetition
label	The label of target sensor, which can be used for sensor and object identification	String	0..*
global_locations	Location of the sensor in the world coordinates		0..1
sensor_id	ID of sensor	String	1
point	Position of the sensor		1
x	X-coordinate	Decimal	1
y	Y-coordinate	Decimal	1
z	Z-coordinate	Decimal	1
unit	Unit of the coordinates	String	1
direction	Direction of the sensor in spherical coordinates		1
use	Senor has direction (Yes/No)	String	1
theta	Angle from X-coordinate in XY plane	Decimal	1
phi	Angle from Z-coordinate	Decimal	1

Table 2 – XML syntax sample of sensor_location_data

```xml
<?xml version="1.0" encoding="UTF-8" ?>
    <sensor_location_data>
        <timecode>DF9F7CB944EF4217</timecode>
        <object>
            <sensor>
                <label>Sensor0</label>
                <global_locations>
                    <sensor_id>Thouch003</sensor_id>
                    <point>
                    <x>392</x>
                    <y>129</y>
                    <z>6405</z>
                    <unit>mm</unit>
                    </point>
                    <direction>
                    <use>Yes</use>
                    <theta>120</theta>
                    <phi>60</phi>
                    </direction>
                </global_locations>
            </sensor>
        </object>
    </sensor_location_data>
```

Bibliography

[b-ITU-T H.430.1]   Recommendation ITU-T H.430.1 (2018), *Requirements for immersive live experience (ILE) services*.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	**Audiovisual and multimedia systems**
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling, and associated measurements and tests
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems