

International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

H.430.1

(08/2018)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS

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

**Requirements for immersive live experience
(ILE) services**

Recommendation ITU-T H.430.1



ITU-T H-SERIES RECOMMENDATIONS
AUDIOVISUAL AND MULTIMEDIA SYSTEMS

CHARACTERISTICS OF VISUAL TELEPHONE SYSTEMS	H.100–H.199
INFRASTRUCTURE OF AUDIOVISUAL SERVICES	
General	H.200–H.219
Transmission multiplexing and synchronization	H.220–H.229
Systems aspects	H.230–H.239
Communication procedures	H.240–H.259
Coding of moving video	H.260–H.279
Related systems aspects	H.280–H.299
Systems and terminal equipment for audiovisual services	H.300–H.349
Directory services architecture for audiovisual and multimedia services	H.350–H.359
Quality of service architecture for audiovisual and multimedia services	H.360–H.369
Telepresence, immersive environments, virtual and extended reality	H.420–H.439
Supplementary services for multimedia	H.450–H.499
MOBILITY AND COLLABORATION PROCEDURES	
Overview of Mobility and Collaboration, definitions, protocols and procedures	H.500–H.509
Mobility for H-Series multimedia systems and services	H.510–H.519
Mobile multimedia collaboration applications and services	H.520–H.529
Security for mobile multimedia systems and services	H.530–H.539
Security for mobile multimedia collaboration applications and services	H.540–H.549
VEHICULAR GATEWAYS AND INTELLIGENT TRANSPORTATION SYSTEMS (ITS)	
Architecture for vehicular gateways	H.550–H.559
Vehicular gateway interfaces	H.560–H.569
BROADBAND, TRIPLE-PLAY AND ADVANCED MULTIMEDIA SERVICES	
Broadband multimedia services over VDSL	H.610–H.619
Advanced multimedia services and applications	H.620–H.629
Ubiquitous sensor network applications and Internet of Things	H.640–H.649
IPTV MULTIMEDIA SERVICES AND APPLICATIONS FOR IPTV	
General aspects	H.700–H.719
IPTV terminal devices	H.720–H.729
IPTV middleware	H.730–H.739
IPTV application event handling	H.740–H.749
IPTV metadata	H.750–H.759
IPTV multimedia application frameworks	H.760–H.769
IPTV service discovery up to consumption	H.770–H.779
Digital Signage	H.780–H.789
E-HEALTH MULTIMEDIA SYSTEMS, SERVICES AND APPLICATIONS	
Personal health systems	H.810–H.819
Interoperability compliance testing of personal health systems (HRN, PAN, LAN, TAN and WAN)	H.820–H.859
Multimedia e-health data exchange services	H.860–H.869
Safe listening	H.870–H.879

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

Recommendation ITU-T H.430.1

Requirements for immersive live experience (ILE) services

Summary

Recommendation ITU-T H.430.1 identifies general requirements and high-level functional requirements for immersive live experience (ILE) services, in order to clarify ILE services. This Recommendation also provides the definition of ILE.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T H.430.1	2018-08-29	16	11.1002/1000/13666

Keywords

Functional requirements, high-level requirements, immersive live experience, ILE, live viewing, public viewing.

* To access the Recommendation, type the URL <http://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID. For example, <http://handle.itu.int/11.1002/1000/11830-en>.

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, 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 TSB patent database at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2019

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

	Page
1 Scope.....	1
2 References.....	1
3 Terms and definitions	1
3.1 Terms defined elsewhere	1
3.2 Terms defined in this Recommendation.....	1
4 Abbreviations and acronyms	1
5 Conventions	1
6 Brief description of the immersive live experience.....	2
7 General requirements of ILE	2
7.1 Displaying real-sized objects.....	2
7.2 Direction of sounds.....	3
7.3 Reconstruction of stage effects.....	3
7.4 Spatial environment.....	3
7.5 Synchronous media representation of multiple assets.....	3
7.6 Augmented information attachment ability.....	3
8 High-level functional requirements of ILE.....	4
8.1 Real-time object extraction.....	4
8.2 Synchronous transmission of multiple media streams	4
8.3 Media processing.....	4
8.4 High-realistic auditory lateralization.....	5
8.5 Video stitching.....	5
Bibliography.....	6

Introduction

In recent times a number of major sporting and music events have not only been broadcast, but have also been delivered to remote sites for public or live viewing with the aim of sharing the emotional experience with spectators in these remote sites as if they themselves were also in the main event venue. At the same time ultra high definition (UHD) broadcasting, known as 4K (resolution of approximately 4000 pixels) and 8K (resolution of approximately 8000 pixels) broadcasting, is becoming increasingly popular. It is therefore to be expected that people all over the world will wish to share in the excitement of viewing sporting or musical events on HD-TV or at public viewing sites in the near future. However the realistic experience offered by flat displays has limitations.

In order to provide high-realistic sensations to audiences at remote sites, immersive live experience (ILE) needs to be implemented to reconstruct event sites virtually with the presentation of real-sized objects and sound direction through the transmission of environmental information together with audio and video streams. For example using ILE musicians could harmonize with player images which are transmitted from several locations. By using ILE, the real objects, such as in this case the musicians at a remote site can collaborate with 3D images of remote musicians which are projected onto a screen. ILE enables virtual concerts with high-realistic sensations.

Implementing ILE needs several technologies such as real-time objects extraction technologies at event sites, spatial location sensing technologies for objects, sound direction identifying technologies, media transport technologies for extracted objects including spatial location information, presentation technologies including 3D projection at remote sites, synchronous technologies with video, sound and lighting, and so on. Although some of these technologies are already established, there are some conditions and/or limitations such as specific content and the pre-arrangement of remote sites. Pre-arrangement of remote sites includes 3D projection mapping and takes a lot of time for adjustment of terminal devices.

In order to share the enthusiasm of event venues with large audiences even if they are in remote sites, far from the event venue, implementing immersive live experience services based on standardized designs is desirable. ILE can help audiences anywhere in the world cheer their favourite sporting teams or musical artists at remote sites even if they are not themselves actually in the event venue and they may feel a sense of togetherness and even become passionate just as though they were actually present in the event venue site.

Recommendation ITU-T H.430.1

Requirements for immersive live experience (ILE) services

1 Scope

This Recommendation identifies the general requirements for immersive live experience (ILE) services. The definition of ILE and its general and high-level functional requirements are also identified. In addition, the requirements are categorized into various parts such as source site, transmission part and presentation sites.

The scope of this Recommendation includes:

- Definition of ILE
- General requirements of ILE
- High-level functional requirements of ILE.

2 References

None.

3 Terms and definitions

3.1 Terms defined elsewhere

None.

3.2 Terms defined in this Recommendation

This Recommendation defines the following term:

3.2.1 immersive live experience (ILE): A shared viewing experience that stimulates emotions within audiences at both the event site and the remote sites, as if the audience at the remote sites had wandered into a substantial event venue and had actually watched the events taking place in front of them. This impression is due to high-realistic sensations provided by a combination of multimedia technologies such as sensorial information acquisition, media processing, media transport, media synchronization and media presentation.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

ILE	Immersive Live Experience
3D	Three Dimensional
MPEG	Moving Picture Experts Group
MMT	MPEG Media Transport
UHD	Ultra High Definition

5 Conventions

In this Recommendation:

- 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.

- The keywords "is recommended" indicate a requirement which is recommended but which is not absolutely required. Thus, this requirement need not be present to claim conformance.
- The keywords "can optionally" indicate an optional requirement that is permissible, without implying any sense of being recommended. This term is not intended to imply that the vendor's implementation must provide the option, and the feature can be optionally enabled by the network operator/service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with this Recommendation.
- The keyword "functions" are defined as a collection of functionalities.
- The keyword "functional block" is defined as a group of functionalities that has not been further subdivided at the level of detail described in this Recommendation.

NOTE – In the future, other groups or other Recommendations may possibly further subdivide these functional blocks.

6 Brief description of the immersive live experience

Immersive live experience (ILE) refers to a shared viewing experience which stimulates emotions within audiences at both the event venue and at remote sites, as if the viewers at the remote sites wandered into a substantial event venue and watched actual events taking place in front of them. This impression is due to high-realistic sensations provided by a combination of multimedia technologies such as sensorial information acquisition, media processing, media transport, media synchronization and media presentation.

7 General requirements of ILE

The aim of ILE services is to reproduce a world where audiences can experience the feeling and excitement of being at an event venue, wherever they are. ILE is a technique for directly transmitting not just the images and sounds of players but also the environment in which the players exist. Realistic sensations can be obtained by displaying real-sized objects such as a human body. For example, if players of sports games are displayed in real-size, audiences might feel a realistic sensation. Regarding the sound of objects, audiences feel a realistic sensation while they recognize the direction of sound. For example, if the voices of players in a sports game come from players displayed on screen, audiences feel it is more realistic.

However current public viewing or live viewing systems cannot deliver nor display real-sized objects because there is no information about the objects and terminal devices on a media stream. Furthermore, it is very difficult to reproduce the direction of sound in terminal devices, since there is no sound direction information.

7.1 Displaying real-sized objects

There are several kinds of display devices which have a wide variety of display sizes. Some display terminal devices have different aspect ratios in pixels and bezels, which relate to the physical outer frames of the display, so the aspect ratio and the bezel size for displaying objects should be considered. In addition, displaying a real-sized object requires the spatial position, for example height from floor and width, of the neighbouring display devices. This spatial position information should also be considered.

High-level requirement 1: Immersive live experience services are recommended to have a capability of displaying real-sized objects on the screens on various types of terminal devices at various positions.

7.2 Direction of sounds

The sounds are produced from objects. The direction of sounds should be aligned with objects which are presented by the display on terminal devices.

High-level requirement 2: Immersive live experience services are required to reproduce the direction of sound on the terminal side.

7.3 Reconstruction of stage effects

Usually, most of public viewing or live viewing sites are different from the source environment or event venue such as a concert venue or sports stadium. One of the major differences is lighting. In order to reconstruct the atmosphere of an event venue at the viewing sites, stage effects or special effects including lighting may increase the high-realistic sensation for viewers. In addition, stage/special effects information might be utilized to emphasize or customize the presentation effect at the viewing sites, for example, spotlights could be added onto the stage at the viewing sites, according to stage/ special effects information.

High-level requirement 3: Immersive live experience services are recommended to reconstruct a suitable atmosphere by using stage effects or special effects including lighting systems.

7.4 Spatial environment

Display devices have a wide variety of specifications, i.e., several kinds of resolutions and screen sizes. In addition, the spatial location of devices which includes horizontal and vertical positions and vent angles is different in each viewing site. In order to reconstruct source objects and environmental situations at the viewing sites, spatial environment information including terminal device capabilities and its spatial location information are required.

High-level requirement 4: Immersive live experience services are required to reconstruct spatial environments according to device capabilities and the spatial location information of terminal devices.

7.5 Synchronous media representation of multiple assets

Most users want to watch sports games in real time and with high-realistic sensations. Real-time content delivery or live video streaming usually requires time synchronization capability between transmitted multiple assets such as video and audio. Even when showing stored content, time synchronization capability is required for high-realistic sensations at the viewing site. For displaying real-sized objects, time synchronization is also required between special environments and displayed objects.

High-level requirement 5: Immersive live experience services are required to use synchronous media representation of multiple assets such as video images with audio information, spatial information and composition information.

7.6 Augmented information attachment ability

Initially most of live viewing usually used to display information such as the current score, red or yellow cards received, time remaining in the sports event, or lyrics of the song in being sung in a concert hall. Augmented information attachment ability can transform these kinds of information into digital content and attach them to a focus zone display such as the centre of the pitch to avoid lack of feeling for the scene.

Secondly, augmented information attachment ability can bring extra information to audiences such as the position or current scoring data displayed just over their favourite players, which people in the real scene cannot obtain at the same time. In addition, more entertainment or

marketing/promotional information will be brought to live broadcasting stakeholders, increasing the market vitality.

High-level requirement 6: Immersive live experience services can optionally have augmented information attachment ability.

8 High-level functional requirements of ILE

In order to realize ILE services including pseudo 3D presentation, several functions such as high-realistic media synchronization are required. This clause describes the minimum set of high-level functional requirements for ILE systems.

8.1 Real-time object extraction

Real-time object extraction from image is a key technology to realize ILE and extracted objects may be displayed as pseudo-3D images at viewing sites. This may be achieved by a combination of contour detection technology using sensor information such as distance, temperature and image processing technology which identifies boundaries of target objects at high-speed and in high-resolution. Real-time object extraction from image can clip purely the target object to be presented as a pseudo-3D image at viewing sites.

High-level requirement 7: Immersive live experience systems are required to extract object information in real time used for displaying high-realistic images at viewing sites.

High-level requirement 8: Immersive live experience systems are recommended to capture spatial information of objects.

8.2 Synchronous transmission of multiple media streams

It is necessary to synchronize several elements such as audio and video streams, lighting, spatial and composition information to reconstruct a virtual field of the event venue at remote sites. One possible technology is MPEG media transport (MMT) which has features to synchronize video and audio streams by using absolute time management. A signalling descriptor of MMT is required to model spatial information such as the size of objects, positional relations and direction of sounds.

High-level requirement 9: Immersive live experience systems are required to transport synchronously video images and audio with a combination of lighting, spatial and composition information.

High-level requirement 10: Immersive live experience systems can optionally store synchronous data including video images and audio with a combination of lighting, spatial and composition information.

8.3 Media processing

For presenting high-realistic images and sounds at viewing sites, media processing needs to handle asset information and to calibrate media related to physical spatial position information such as the size of the screen, resolution of displays and spatial position. Asset information is required to process aligning viewing sites conditions by using the spatial size of images and auditory lateralization information, so that the virtual field of the event venue is reconstructed with high-reality at the viewing site, depending on the specifications of each viewing site.

High-level requirement 11: Immersive live experience systems are required to process several media such as audio and video stream, lighting and spatial information, in order to reconstruct virtual field in viewing sites.

8.4 High-realistic auditory lateralization

High-realistic auditory lateralization aims to reconstruct a virtual acoustic field into a wide viewing area. This enables reproduction of precise sound direction from objects projected on a wide screen, thus audiences can feel the voices and sounds as if the projected objects emitted the sounds.

High-level requirement 12: Immersive live experience systems are recommended to have high-realistic auditory lateralization for reproducing precise sound direction at viewing sites.

8.5 Video stitching

The video stitching technology creates extra-wide angle video images from multiple cameras and adjusts images in order to reconstruct high-realistic field images in real time at viewing sites. This technology is usually used for very large sports venues and widely dispersed fields such as in athletics events.

High-level requirement 13: Immersive live experience systems can optionally have video stitching functions, if several cameras are used at the event site.

Bibliography

- [b-ITU-T H.430.2] Recommendation ITU-T H.430.2 (2018), *Architectural framework for immersive live experience (ILE) services*.
- [b-ITU-T H.430.3] Recommendation ITU-T H.430.3 (2018), *Service scenario of immersive live experience (ILE)*.

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