

International Standards on Communication Technology for Immersive Live Experience Adopted by ITU-T

Jiro Nagao

Abstract

International standards on immersive live experience (ILE) have been adopted by the International Telecommunication Union - Telecommunication Standardization Sector in November 2019. ILE captures an event and transports it to remote viewing sites in real time and reproduces it with high realism. It is expected that the series of standards on ILE will facilitate the creation of a world where people can enjoy highly realistic reproduction of events in real time wherever they are. This article is an introduction to such ILE standards.

Keywords: immersive live experience, live public screening, highly realistic real-time reproduction

1. Immersive live experience

Public live screening, where popular events such as music concerts and sports games are relayed to movie theaters and large screens outside the actual event sites and people at remote locations can enjoy the events at the same time as the actual events, has become increasingly popular. The advantage of public live screening is that those who cannot go to the actual events for various reasons, such as they live far from the actual event sites, can enjoy the event at more convenient places and share their emotions and experiences with each other.

Public live screening, however, has the drawback of being less realistic than the actual event. Therefore, technologies to enable more realistic public live screening is needed. Such technologies may enable images of artists in an actual event site to be shown at viewing sites with high realism, as if they were performing in front of the remote audience, sounds to come from the actual objects on the screen, or the atmosphere, i.e., the vibration or heat from the actual event site, to be reproduced. If such technologies are

feasible, public live screening with much higher realism can be achieved. The feeling as if one is at the actual event site from images (videos) and sounds is called *immersive experience*, and there have been studies on achieving such an experience. When immersive experience is achieved in real time, it is called *immersive live experience (ILE)*. ILE enables real-time experience of remote events with highly realistic sensations anywhere in the world. It is especially advantageous in cases such as sports matches in which simultaneity is an important factor of entertainment.

ILE requires technologies to capture various conditions (environment information) of the event in detail, transport them along with video and audio to viewing sites at remote locations, and reproduce them in real time. The information to be captured includes three-dimensional (3D) positions of persons and objects, positions of sound sources, and lighting-control signals. Reduction in the conversion time of formats between different systems and interconnectivity of content transport to remote locations are necessary for real-time reproduction. Therefore, international

Table 1. ITU-T standards on ILE.

Number	Title	Adopted
H.430.1	Requirements for immersive live experience (ILE) services	Aug. 2018
H.430.2	Architectural framework for immersive live experience (ILE) services	Aug. 2018
H.430.3	Service scenario of immersive live experience (ILE)	Aug. 2018
H.430.4	Service configuration, media transport protocols, signalling information of MPEG media transport for immersive live experience (ILE) systems	Nov. 2019

Table 2. Requirements of ILE.

No.	Requirements	Required/Recommended/Optional
1	Life-size display	Recommended
2	Direction of sound	Required
3	Atmosphere reconstruction	Recommended
4	Spatial reconstruction	Required
5	Synchronous representation	Required
6	Augmented information	Optional
7	Real-time object extraction	Required
8	Spatial object information	Recommended
9	Synchronous transmission	Required
10	Data storage	Optional
11	Reconstruction processing	Required
12	Auditory lateralization	Recommended
13	Video stitching	Optional

standards to ensure global interconnectivity are needed. However, there have not been standards specialized in transporting environment information in real time along with video and audio and reproducing an event at viewing sites with high realism.

2. International standards on ILE

The International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) Study Group (SG) 16^{*1} launched Question 8 “Immersive live experience systems and services” in 2016 to address the issues regarding the implementation of ILE that provides highly realistic sensations to many audiences anywhere in the world in real time by internationally connecting the systems at actual event sites (source sites) and remote viewing sites. The ITU-T Recommendation H.430 series has been adopted as the result of the activities of this Question [1–4]. **Table 1** shows the ITU-T standards on ILE.

2.1 H.430.1

Standard H.430.1 defines ILE and identifies its requirements. **Table 2** shows these requirements, which include life-size display of images and auditory lateralization^{*2} for highly realistic reproduction of events, real-time extraction to reproduce images of people at the viewing sites, and real-time synchronized transport of environment information along with video and audio.

2.2 H.430.2

Standard H.430.2 identifies the high-level functional architecture and the general model of ILE. It also introduces candidate technologies for ILE functions. **Figure 1** shows the high-level functional

*1 ITU-T SG16: The ITU is the United Nations specialized agency in the field of telecommunications, information and communication technologies. The ITU-T is responsible for issuing Recommendations on this field. SG16 leads ITU’s standardization work on multimedia coding, systems, and applications.

*2 Auditory lateralization: To reproduce or the methods of reproducing direction of sound using multiple loud speakers.

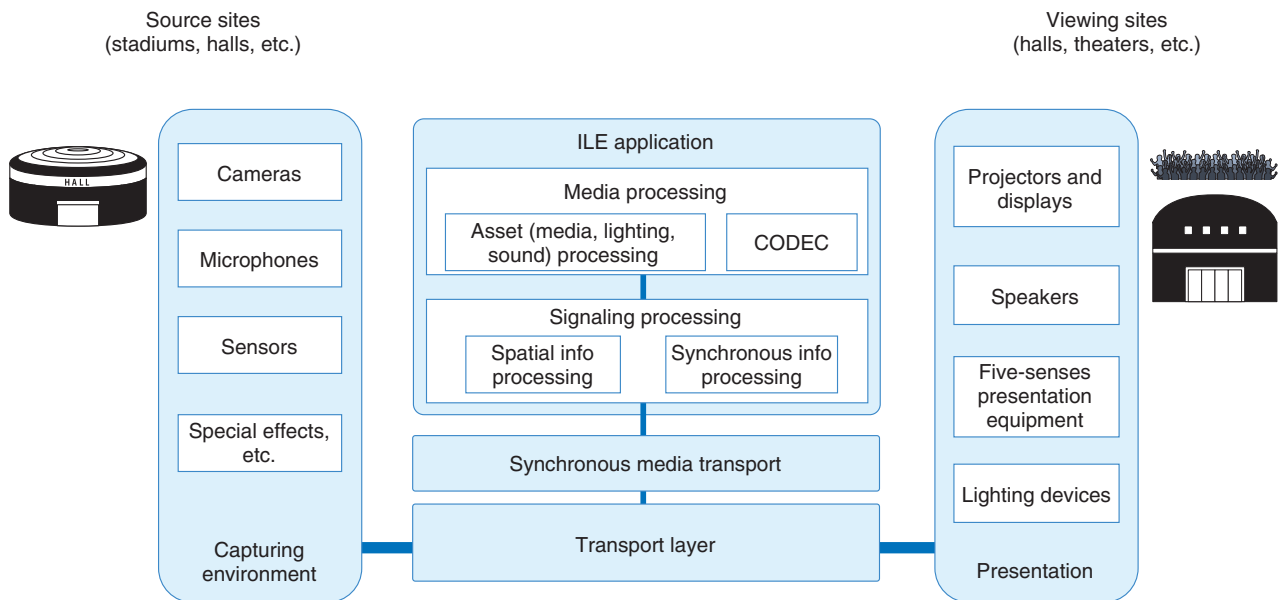


Fig. 1. High-level architecture of ILE system.

architecture of ILE. The architecture consists of the following functionalities.

(1) Capturing environment

Environment information, such as the video, audio, positions of people, and lighting-control signals at the source site, is captured with cameras, microphones, sensors, etc.

(2) Synchronous media transport

Synchronous media transport is a function that transports multiple media synchronously.

(3) Transport layer

Media data at the source site are transported to the remote sites with considerations on content protection and network delay.

(4) ILE application

Asset processing performs, for example, extraction of images of people from video and integration of videos to achieve a higher sense of realism. CODEC encodes and decodes media content. Spatial information processing integrates, for example, the 3D position information obtained from camera images and positional sensors. Synchronous information processing synchronizes content, such as video, audio, lighting control, and sensor information, for reproduction with the presentation functionality.

(5) Presentation

The presentation functionality displays video, lateralizes audio, reproduces lighting, etc. at the viewing sites. Synchronized presentation of such media pro-

vides highly realistic sensations to the audience.

2.3 H.430.3

Standard H.430.3 categorizes ILE service scenarios with use case examples, providing clear ideas of novel experiences enabled by ILE services. Some of the examples are as follows:

(1) Live broadcast service of first-person synchronous view

The first-person view of an athlete in a fast-moving vehicle such as a racing car or bobsled is transported and reproduced at viewing sites in real time. Additional information, such as speed, acceleration, current ranking, and remaining distance, can be presented to enhance the excitement of the competition.

(2) On-stage holographic performance service

A live concert performance is captured with ultra-high definition cameras and high-quality microphones. The captured data are processed and transported to the viewing sites then used to reproduce the event in 3D or pseudo-3D images with high-quality audio and special effects, providing the audience with highly realistic sensations.

(3) Multi-angle viewing service

The images of artists or athletes are captured with cameras from multiple directions and transported to remote viewing sites in real time along with high-quality audio and environment information such as the 3D positions of the artists or players. The data are

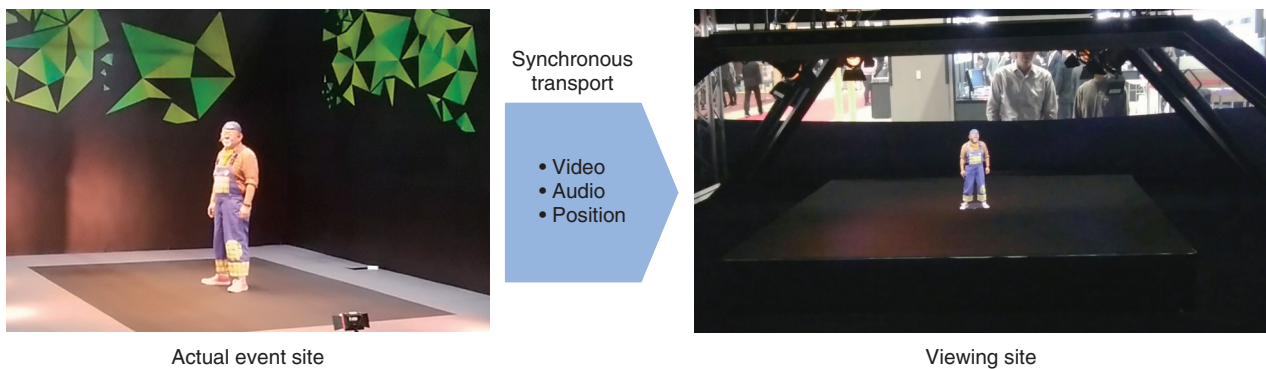


Fig. 2. A prototype system for the multi-angle viewing service.

reproduced at the viewing sites on special display devices with multiple displays to show views from all directions. The 3D position information is used to add a stronger sense of depth to the displayed images. Auditory lateralization can be provided with multiple loud speakers. A prototype system built for this purpose is shown in **Fig. 2**. The display device can show not only the view from front but also views from sides and behind, providing the impression as if moving around the actual event site.

2.4 H.430.4

Standard H.430.4 identifies the MPEG Media Transport (MMT^{*3}) profile for ILE to provide synchronous transport of spatial information, such as the 3D position of objects, along with video and audio for reproduction of events with high realism.

The spatial information includes the following:

- Size (width, height, depth) of the source site
- Equipment information (equipment type, model, location, direction, size) of the source site
- 3D positions, positions in video, and bounding rectangles of the performers on stage

*3 MMT: A standard for media transport developed by the international standards developing organization ISO/IEC JTC 1/SC 29/WG 11 (International Organization for Standardization/International Electrotechnical Commission Joint Technical Committee 1/Subcommittee 29/Working Group 11), also known as MPEG (Moving Picture Experts Group).

Spatial information enables the reconstruction of 3D positions of the performers at the viewing sites while adapting to the difference between the size and location of equipment at the source and viewing sites.

3. Future prospect

Further development of Recommendations, including reference models and implementation guidelines of the presentation environment, and incorporating related technologies such as virtual reality are planned by ITU-T in collaboration with other standards developing organizations.

References

- [1] H.430.1: Requirements for immersive live experience (ILE) services, <https://www.itu.int/rec/T-REC-H.430.1/en>
- [2] H.430.2: Architectural framework for immersive live experience (ILE) services, <https://www.itu.int/rec/T-REC-H.430.2/en>
- [3] H.430.3: Service scenario of immersive live experience (ILE), <https://www.itu.int/rec/T-REC-H.430.3/en>
- [4] H.430.4: Service configuration, media transport protocols, signalling information of MPEG media transport for immersive live experience (ILE) systems, <https://www.itu.int/rec/T-REC-H.430.4/en>



Jiro Nagao

Research Engineer, NTT Service Evolution Laboratories.

He received a Ph.D. in information science from Nagoya University, Aichi, in 2007. He joined NTT the same year. He was a technical leader of commercial video streaming services at NTT Communications and is currently engaged in research on immersive media and presentation at NTT Service Evolution Laboratories. Since 2019, he has contributed to the international standardization efforts on immersive live experience (ILE) of ITU-T SG16.