International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

H.430.5

(08/2020)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS

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

Reference models for immersive live experience (ILE) presentation environments

Recommendation ITU-T H.430.5

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.430.5

Reference models for immersive live experience (ILE) presentation environments

Summary

In order to reproduce a highly realistic experience and sense of immersiveness for audiences at a viewing site on par with the experience of audiences at the event site at the same time, and also to reduce the design and setup times of immersive live experience (ILE) viewing sites, Recommendation ITU-T H.430.5 specifies reference models for the immersive live experience (ILE) presentation environments. There are three basic types of ILE presentation environment: a proscenium style that has a one-sided stage or screens, an open style that has island stages or runways from a single-sided stage, and an arena style presentation environment that has one island stage. This Recommendation provides one reference model each for the proscenium, open, and arena style presentation environments. This Recommendation also provides functional blocks and implementation guidelines for ILE viewing sites as informative information.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T H.430.5	2020-08-13	16	11.1002/1000/14338

Keywords

Functional blocks, ILE, implementation guidelines for ILE viewing sites, presentation environment, reference model.

* 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/.

Table of Contents

Introduction

The most significant feature of the immersive live experience (ILE) is that audiences at viewing sites can enjoy a highly realistic experience and sense of immersiveness on par with the experience of audiences at the event site. To enable this feature, it is necessary to set up the viewing sites with elements such as projection mapping, audio systems suitable for each viewing site to achieve auditory lateralization, and other stage effects such as lighting, vibration, and smoke for reproducing the event in a way that stimulates the five senses. However, viewing sites generally consist of halls that are of completely different sizes (e.g., height and width) and shapes (e.g., rectangular or circular), that have a different numbers of seats, and have different stage types (e.g., theatre and island types). Therefore, a highly realistic atmosphere at the viewing sites needs to be reproduced by considering the unique features of the individual sites. The presentation environments for ILE services are currently constructed site by site. This takes a very long time and comes at huge labour and financial costs.

Implementation guidelines for the ILE presentation environment based on a reference model of the viewing site with related parameters may help to reduce the time and costs. Such guidelines would greatly help with the efficient implementation of the presentation environment.

This Recommendation provides three reference models for the ILE presentation environment. The functional blocks and implementation guidelines are attached in the appendices.

Recommendation ITU-T H.430.5

Reference models for immersive live experience (ILE) presentation environments

1 Scope

In order to reproduce highly realistic sensations or immersiveness on par with the experience of the audience at the event site, and to reduce the design and set-up time of immersive live experience (ILE) viewing sites, this Recommendation provides three reference models for the ILE presentation environment. This Recommendation supports the ILE services as described in the ITU-T H.430-series Recommendations. It also introduces an example of functional blocks for the presentation environment as well as implementation guidelines based on the reference models.

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

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

[ITU-T H.430.3] Recommendation ITU-T H.430.3 (2018), *Service scenario of immersive live experience (ILE)*.

[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*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following term defined elsewhere:

3.1.1 immersive live experience (ILE) [ITU-T H.430.1]: 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.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

HDMI High-Definition Multimedia Interface

ILE Immersive Live Experience

LCD Liquid Crystal Display

MMT MPEG Media Transport

USB Universal Serial Bus

WFS Wave Field Synthesis

5 Conventions

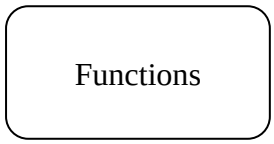
In this Recommendation:

– 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.

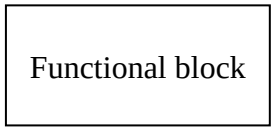
– The keywords "is recommended" indicate a requirement which is recommended but which is not absolutely required. Thus, this requirement need not be present to claim conformance.

– 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. It is represented by the following symbol in this Recommendation:



– 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. It is represented by the following symbol in this Recommendation:



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

Frame borders of "functions" and "functional blocks", and relational lines among "functions" and "functional blocks" are drawn with solid lines or dashed lines. The solid lines mean required functionalities or relations. On the other hand, the dashed lines mean optional functionalities or relations.

6 Overview of ILE viewing sites

6.1 Types of presentation environments

As defined in [ITU-T H.430.2], there are several types of presentation environments, such as proscenium style, open style, and amphitheatre (arena style). These types are illustrated in Figure 6-1.

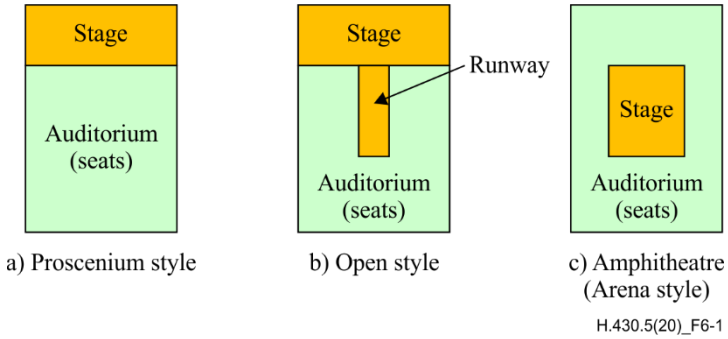


Figure 6-1 – Types of presentation environments

Figure adapted from [ITU-T H.430.2]

a) Proscenium style

In most theatres, audiences view a stage or screens in front of them. As shown in Figure 6-1(a), the stage is located at one end of theatre, and the seats for the audience, called the auditorium, are located at the other end. This type of theatre is called proscenium style. The details are discussed in [ITU-T H.430.2].

b) Open style

Most concert halls have one or more stages and/or runways, as shown in Figure 6-1(b), so that audiences can view the performance from different angles. This type of theatre is called open style. The open style has many variations, such as a theatre that has a stage with one or more runways, or a theatre that has a stage with one or more island type stages. The details are discussed in [ITU-T H.430.2].

c) Amphitheatre (Arena style)

Amphitheatres (arena style), as shown in Figure 6-1(c), have one field or one stage at the centre of the theatre, and audiences may be seated around the stage, which means they can watch performances from a 360-degree angle. The details are discussed in [ITU-T H.430.2].

6.2 Proscenium style presentation environment

In the proscenium style, the entire audience is seated in one direction facing the stage. In order to reconstruct this environment at viewing sites, one or more display(s) are placed on a single side. One example is a synchronous live talk show or a lecture presentation, where the feed is transmitted by live streaming from the event site to one or more viewing sites.

There are three required functionalities at the viewing site side: controlling, receiving, and displaying. For providing ILE services in the proscenium style presentation environment, the following functions are required.

– Controlling functionality:
  – Controlling:

    Hardware and the operating system (OS) are the controlling functions.

- Resident:

  The resident function provides a user interface such as a menu screen that enables content selection. It might have a function for the setup screen of display devices. The resident function calls a renderer in order to play content.

- Input interface:

  The input interface connects the input devices to allow users to control or operate the receiving functionality. A universal serial bus (USB) might be used for this interface.

- Input devices:

  The input devices are usually keyboards and pointing devices.

- Receiving functionality:
  - Communication interface:

    Communication interfaces are used to handle data of user interfaces, coded content data, and other controlling information. For example, HTTP, HTTPS, and MPEG media transport (MMT) can be used as the communication interfaces.

  - Receiving function for streaming content:

    The receiving function for streaming content can receive content sent from the event site via the communication interface and immediately transfer the received streaming content to the displaying functionality.

  - Receiving function for downloaded content:

    The receiving function for downloaded content can receive content sent from the event site via the communication interface and then store the content before later transferring it to the displaying functionality.

  - Communication devices:

    The communication devices are used for connecting networks through the communication interface. Network interface cards can be used as the communication device.

- Displaying functionality:
  - Displaying function:

    The displaying function is a set of core technologies consisting of displaying user interfaces and a renderer for displaying content. This function treats not only video and audio but also various signals for stage effects such as lighting information.

  - Renderer:

    This function decodes various types of video and audio content into video and audio signals. It may also decode other information such as lighting information.

  - Displaying interface:

    The displaying interface is an interface for video content between the receiving function and the display devices. High-definition multimedia interface (HDMI) might be used for this interface.

  - Display devices:

    The display devices output video content. Liquid crystal displays (LCDs), plasma displays, and projectors are generally used.

6.3 Open style presentation environment

The open style presentation environment is similar to the proscenium style but has stages at the centre and/or (both) sides of the viewing site. Some ILE services, such as those used for Kabuki[1] plays and musical concerts, can be considered to use the open style presentation environment by combining real persons (actors or players) with pseudo-3D images (objects) displayed on screens. There are several styles, such as one centre stage, two side stages, and front and rear stages. On the stages, real persons may perform simultaneously with images displayed on movable screens.

The required functionalities of the open style presentation environment are the functions for combining real persons and images in addition to all functionalities of the proscenium style presentation.

6.4 Arena style presentation environment

In the arena style presentation environment, the audience is seated in a circle around the stage, that is, to the left, right, front, and back of the stage. In order to reconstruct this environment at the viewing sites, at least four displays are required. As an example, a four-surface display device is shown in Figure 6-2. In this device, the events are reproduced at the viewing site at a fraction of the actual size at the event venue. Four half-mirrors (marked 1 to 4 in the figure) are combined in a four-sided pyramid shape, and each half-mirror displays a virtual image by reflecting the images of the flat display device located above the half-mirror. Each half-mirror is installed at a 45-degree tilt to display virtual images in a vertical fashion, and a floor (or stage) is installed inside the half-mirrors to function as the floor of the virtual images.

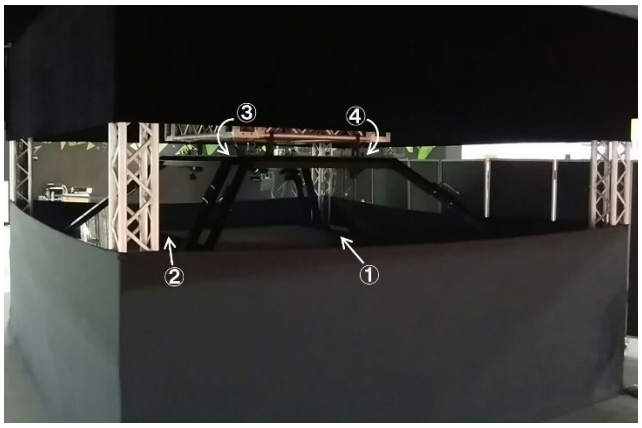


Figure 6-2 – An example of an arena style presentation environment

Each of the images displayed on the half-mirrors looks as if it is floating in the empty space above the stage inside the pyramid-shaped display device.

By crossing the virtual image planes of the four half-mirrors, and by adjusting the distance between the half-mirrors and the related display devices, the virtual images reflected by the half-mirrors can be located at the centre of the stage. It is required to display images synchronously on the four display devices. This technique enables virtual images to be displayed at the centre of the stage when viewed from any of the four directions around the arena style presentation environment.

Several functions are required in order to arouse the sense of depth by the virtual images displayed on the four-surface display device. A depth expression function, which performs the image processing to add depth expression to the images, is required. In addition, a position measurement

[1] Kabuki is a traditional Japanese popular dance-drama with roots tracing back to the 17th century. It is a UNESCO Intangible Cultural Heritage.

function is required to collect the three-dimensional position information of the objects. At the event site, it is necessary to measure and track the three-dimensional position of the objects. It is also required to integrate the video and the various kinds of information collected by object extraction, position measurement, and object tracking. Synchronous transmission of the video, audio, and information from the other source site to the viewing sites is also required.

The controlling, receiving, and displaying functionalities of the arena style presentation environment are the same as those of the proscenium style presentation environment.

7 Reference models for ILE presentation environments

7.1 Reference model for the proscenium style presentation environment

The proscenium style presentation environment is essentially a basic style in which many audiences are simultaneously provided with a single-sided view. The reference model of the proscenium style presentation environment consists of the location of displays, object tracking on the environment, auditory lateralization of displayed objects, and other consideration points.

7.1.1 Construction of proscenium style reference model

The location of displays is shown in Figure 7-1. There are two displays, one each for the foreground and background screens, and two speakers for providing stereo audio. In order to reproduce 3D images, if a screen is constructed from several displays, the x-y-z coordinates of the display positions need to be measured. The foreground screen is usually transparent so as to transmit images displayed on the background screen.

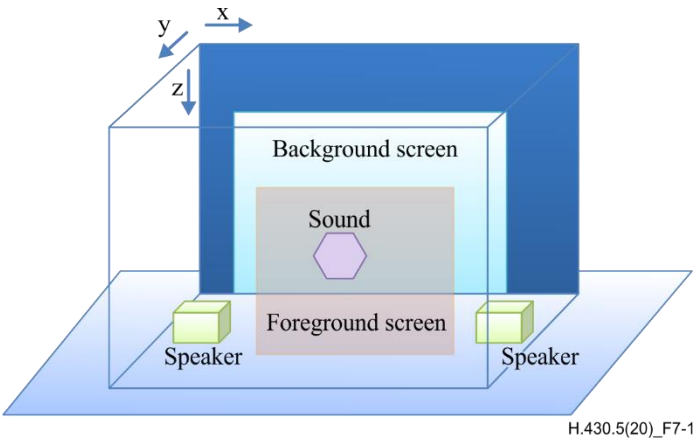


Figure 7-1 – Location of displays in proscenium style reference model

One of the simplest techniques for showing a pseudo-3D image on a transparent display as the foreground screen is Pepper's ghost using a half-mirror. A half-mirror tilted at a 45-degree angle reflects the image on a flat display device to provide an aerial virtual image to the audience. Figure 7-2 shows the mechanism of the aerial virtual image, which has one display and one half-mirror.

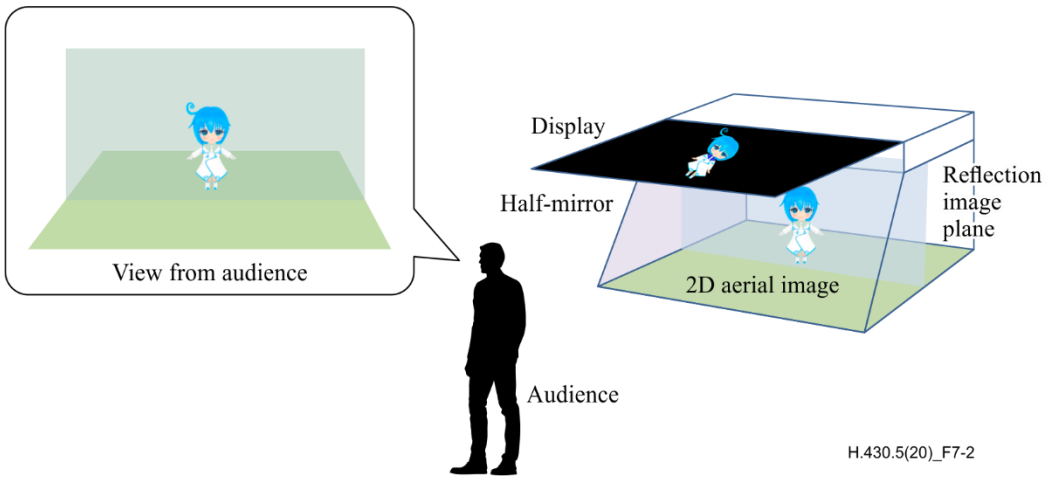


Figure 7-2 – Mechanism of displaying a virtual image by reflection

As shown in Figure 7-2, audiences feel as if they are looking at a real image, when actually it is a virtual image floating in the air behind the half-mirror, which is a reflection of the display on the half-mirror. The principle behind this mechanism is called Pepper's ghost, which is an old illusion technique dating back to the 1860s. The displayed image on the half-mirror can be projected in a pseudo-3D manner at the centre of the floor by arranging distances between the half-mirror and the associated display.

7.1.2 Options for the reference model

The above reference model is a basic construction model. There are several options to expand it and these options are numbered 1 to 5 as follows:

Basic model – Foreground and background screens

1) Extracted images at the event site are displayed on the foreground screen, and the background image is displayed on the background screen at the same time. The foreground screen can project pseudo-3D images by using the Pepper's ghost principle.

Display options

1) Only foreground images

   This option uses one or more foreground image(s) without any background image. This option is mainly used for lecture style events in which only the image of the presenter is provided. This option may use additional images such as presentation screens or slides.

2) Only background images

   This option uses one or more background screen(s) without any foreground images. The background screen should have a very high definition resolution (such as 8K or higher) and provide very wide views so that audiences can get a highly realistic sensation.

Physical object options

3) Static physical objects

   In addition to the basic model, one or more static physical object(s) can be located between the foreground and background displays or in front of the foreground display so that audiences feel a more realistic sensation. Examples of static physical objects include a table-tennis table, a badminton net, and a lecture podium.

4) Moving physical objects (real performers)

In addition to the basic model or the static physical objects option, there can be moving physical objects. In particular, one or more real persons can perform with the displayed images at the viewing sites so that the audience can enjoy a new type of entertainment.

Audio option

5) Auditory lateralization

In addition to the above options, auditory lateralization can give the audience a more realistic sensation, especially when a very wide screen is used. In contrast to stereo sound, which provides sound from the left and right directions, auditory lateralization provides 3D positions of sounds that align with the displayed pseudo-3D images.

The options above are classified into three groups – display, physical objects, and audio and can be combined with each other, such as when combining "static physical objects" with "auditory lateralization" for a relatively big hall with a big screen.

7.2 Reference model for the open style presentation environment

The open style presentation environment is widely used in concert halls, which typically have one or more runways and a one-sided front stage. The front stage might be a proscenium style presentation environment, or a normal stage that contains no immersive images. The reference model can be based on the proscenium style with some additional runway stages.

The reference model for the open style presentation environment consists of the location of displays and other options.

7.2.1 Construction of open style reference model

An overview of the reference model is shown in Figure 7-3. It consists of a one-sided front stage and one centre runway stage. The front stage might be a proscenium style presentation environment, and it could also be a physical stage for performing physical objects (actors or musicians) or flat displays that show videos.

On the runway stages, there are several transparent screens for image projection, such as transparent displays, transparent reflectors, and scrims. The transparent screen might be set up the same way as the proscenium style presentation environment. If the audience is seated on both sides of the runway stage, it is required to use double-sided displays that have two transparent screens for both sides. The audio environment is the same as the proscenium style.

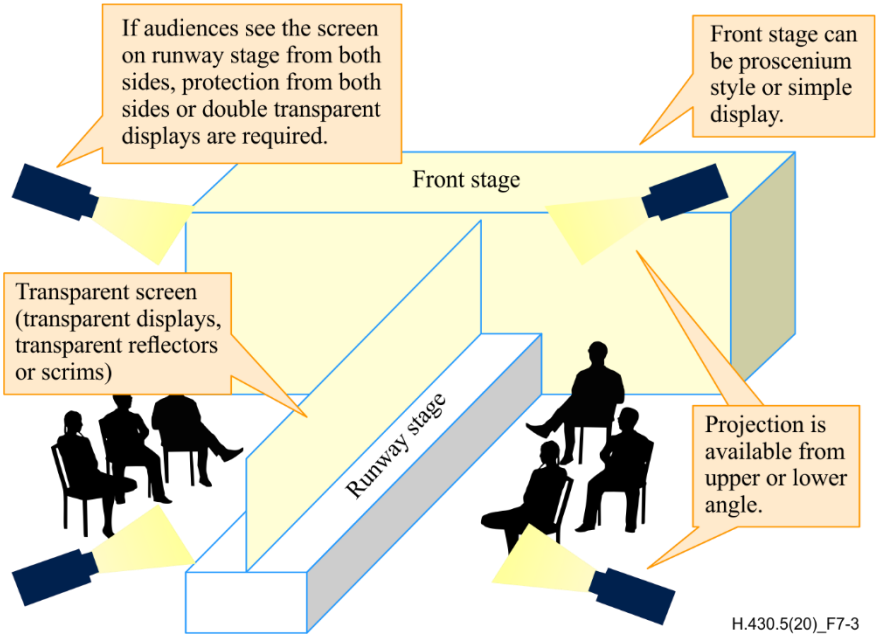


Figure 7-3 – Location of displays in open style reference model

7.2.2 Options for the reference model

The above reference model is the basic construction of the open style presentation environment. There are several options for expanding it and these options are numbered 1 to 10 as follows.

Runway options

1) One runway stage in the middle of the hall

   One runway is located in the middle of the hall. This option is typically used for fashion shows and for most musical concerts.

2) One runway stage at the right or left side of the hall

   One runway is located on either (leftmost or rightmost) side of the front stage.

   This option is usually used for Kabuki plays.

3) Two runway stages on both sides of the hall

   Two runways are located at both ends of the front stage.

4) Combination of above options

   One or more runways are located at arbitrary positions. This option is used in relatively large venues.

Display options on runway stage

5) Fixed displays

   This option uses one or more fixed displays on the runway stage. It includes the proscenium style ILE presentation environment.

6) Moving displays

   This option uses one or more moving displays on the runway stage. It is implemented by means of a movable transparent display that can project the image of a performer. In this case, it is better to use double-sided displays so that front side of the transparent display projects the front view of the performer, and the back side projects the back view.

Display options on front stage

7) Simple displays

This option uses one or more fixed displays on the front stage. It is the simplest option for the front stage. The displays project background images of the performance.

8) Physical stage

This option uses one or more physical objects including real persons on the front stage. This option can be used for making the scene more immersive to the audience. One example of a physical object is a DJ booth with a DJ in a musical concert.

9) Proscenium style ILE presentation environment

This option uses the proscenium style ILE presentation environment on the front stage. It requires a relatively large venue.

Audio option

10) Auditory lateralization

In addition to the above options, auditory lateralization makes audiences feel a more realistic sensation, especially when a very wide screen is used. In contrast to stereo sound, which provides audio from the left and right directions, auditory lateralization provides 3D positions of sounds that align with the displayed images and display locations.

The runway option is the option for physical ILE viewing sites. From the implementation point of view, display types used for runway stages are categorized into two types: a double-sided display, which consists of two displays arranged back to back, and a single-sided display. The options above are classified into four groups: display types on runway stages, a display option on runway stages, a display option on the front stage, and an audio option. The options can be combined with each other, such as combining "moving displays" and "auditory lateralization" for a relatively big hall with a big screen.

7.3 Reference model for the arena style presentation environment

The arena style presentation environment can provide multi-angle viewing for several audiences at the same time. The reference model for the arena style presentation environment consists of the location and number of displays, object tracking on the environment and other consideration points.

7.3.1 Construction of arena style reference model

The arena style presentation environment has four screens, which encompass one floor. Each screen is the same as the foreground screen of the proscenium style presentation environment described in clause 7.1.

Figure 7-4 shows the reference model for the arena style presentation environment using the Pepper's ghost technique. There can be more than two speakers located on each side.

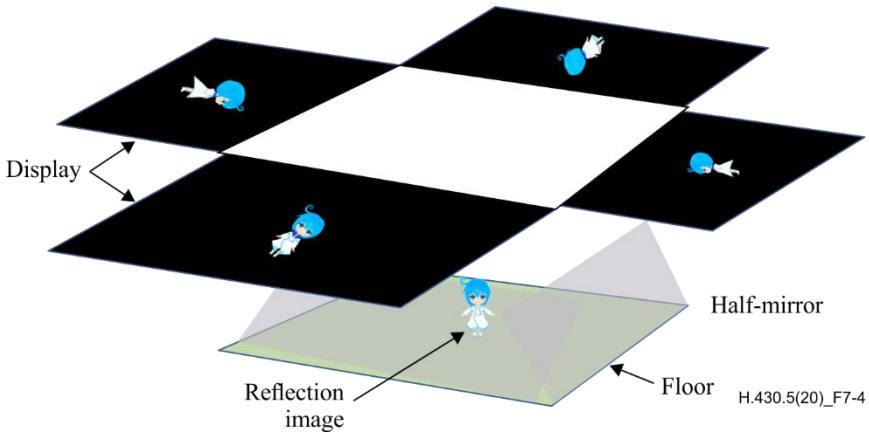


Figure 7-4 – Reference model for arena style presentation environment

In this reference model, four displays are located above the model and four half-mirrors are constructed in a quadrangular pyramid shape. There is a physical stage as a floor inside the four half-mirrors. Half-mirrors are tilted at a 45-degree angle to reflect images of flat display devices and provide virtual images for the audience. The mechanism of displaying the aerial virtual image on each side of the model is the same as that of the proscenium style presentation environment described in clause 7.1.

Displayed images on half-mirrors can be projected in a pseudo-3D manner at the centre of the floor by arranging the distances between a half-mirror and associated display and between the half-mirrors themselves. Due to synchronously displayed images from four angles, audiences around the arena style presentation device can view the pseudo-3D virtual object on the floor from any angle.

In the arena style presentation, four display devices project images from each angle, so cameras need to be set up in four directions around the target objects at the event site: front, rear, right, and left. Object images need to be extracted from whole images captured by each camera in real time. In order to give audiences at the viewing sites the feeling of depth, 3D depth information of the object needs to be captured at the event site, and images are then processed for depth perception and presentation on the basis of the depth information. Therefore, the key functions of the arena style presentation environment are depth perception and presentation. Extracted object images with integrated positional information need to be presented on four-sided presentation display devices.

7.3.2 Options for the reference model

The arena style presentation environment has several options and these options are numbered 1 to 3 as follows:

Physical object options

1) Static physical objects

   Similar to the proscenium style presentation environment, one or more static physical objects can be placed on the stage, so that audiences feel a more realistic sensation. Examples of static physical objects include a table-tennis table, and a badminton net.

2) Moving physical objects (real performers)

   In addition to the static physical objects option, there can be moving physical objects. In particular, one or more real persons can perform with the displayed images at the viewing sites so that the audience can enjoy a new type of entertainment.

Audio option

3) Auditory lateralization

In addition to the above options, auditory lateralization makes audiences feel a more realistic sensation, especially when a very wide screen is used. In contrast to stereo sound, which provides audio from the left and right directions, auditory lateralization provides 3D positions of sounds that align with the displayed images and display locations.

The options above are classified into three groups – image expression, physical object, and audio – and can be combined with each other, such as combining "static physical objects" and "auditory lateralization" for a relatively big hall with a big screen.

Appendix I

Example of functional blocks for presentation environment

(This appendix does not form an integral part of this Recommendation.)

One example of functional blocks described in this Appendix may be used for all three types of presentation environment. For the open style and arena style presentation environments, one or more image configurators might be used depending on the number of screens. In addition, the number of other functional blocks might depend on the types and the number of facilities of the actual viewing sites.

Information from the event site is transmitted on MMT signals to viewing sites via an immersive live experience application. The MMT signals received at the viewing sites are processed by the following five functional blocks, as shown in Figure I.1.

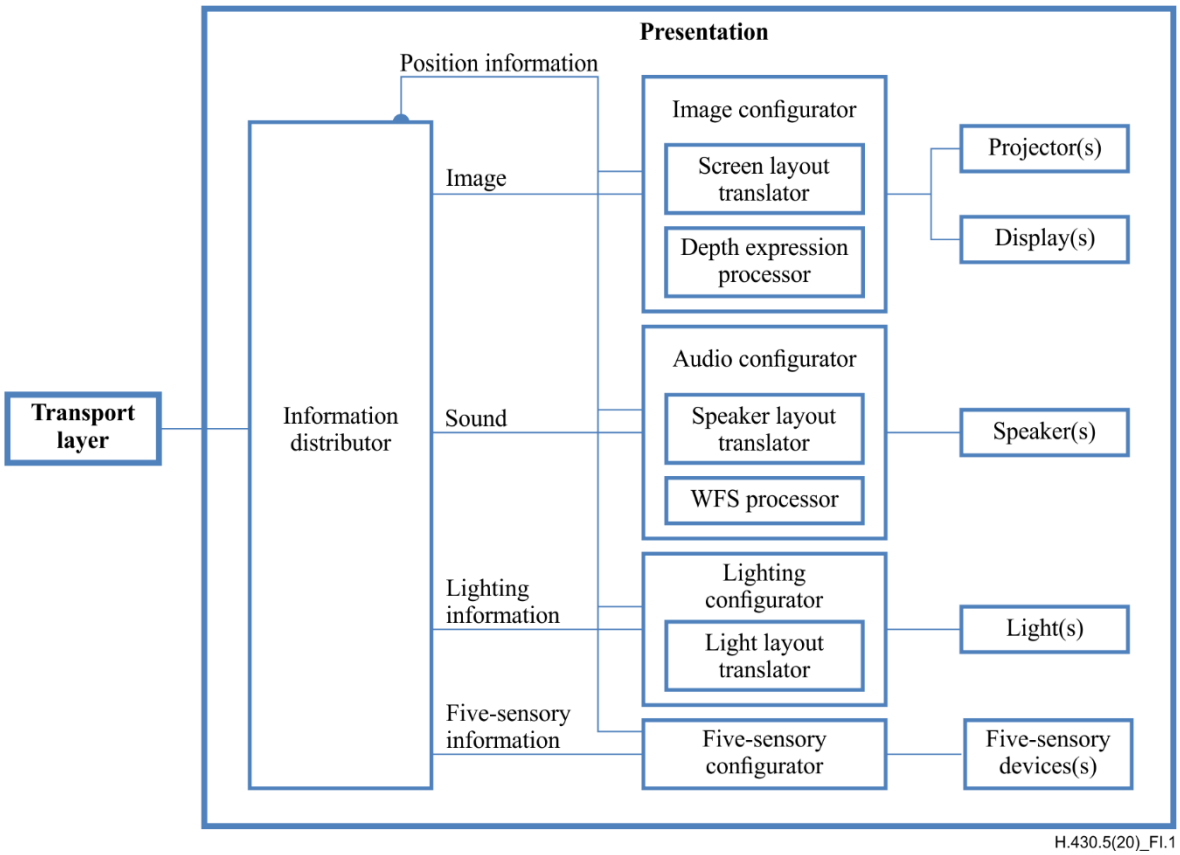


Figure I.1 – Five functional blocks for media processing at viewing sites

1) Information distributor

The information distributor receives the data from the transport layer and divides it into image, audio, lighting, five-sensory information, and position information and then distributes the data to the relevant postprocessors.

2) Image configurator

The image configurator is a function that processes the images received by the presentation functionality in accordance with the conditions unique to the particular presentation environment, such as display types and the number of screens at the viewing site. The image configurator consists of a screen layout translator and a depth expression processor in order to project many types of images, including 3D images.

Screen layout translator

The screen layout translator adjusts the images to the particular displaying equipment of the presentation environment. This adjustment may include selecting displays or projectors to show the images and changing and cropping the images to fit the displays or projectors in a way that best reproduces the event using the equipment available at the viewing site. The screen layout translator assigns the screens (displays or projectors) to images that best reconstruct the scene. This assignment may be determined according to the screen layout at the viewing site and the position information of the screens, cameras, objects, etc. at the source site.

Depth expression processor

The depth expression processor transforms the images and adjusts the display devices to give a realistic sense of depth to the images in accordance with the 3D position of the object in the images. The transformation and adjustment may include changing the shape (e.g., perspective distortion), adjusting the image position, and adjusting the brightness.

3) Audio configurator

The audio configurator is a function that provides audio lateralization to the audience. It processes the audio signals received by the presentation functionality in accordance with the conditions unique to the particular presentation environment at the viewing site. The audio configurator function consists of two functions: a speaker layout translator and a wave field synthesis (WFS) processor.

Speaker layout translator

The speaker layout translator assigns the speakers to audio signals that best reconstruct the sound at the source site. This assignment may be determined according to the speaker layout at the viewing site and the position information of the speakers, mics, objects, etc. at the source site. The speaker layout translator may also perform upward and downward conversion of the channels.

WFS processor

The WFS processor performs wave field synthesis (WFS) to reconstruct the sound field of the source site at the viewing site for highly realistic audio lateralization. It uses the audio signals, the 3D position information of the objects (sound source), and the particular speaker layout information at the viewing site.

4) Lighting configurator

The lighting configurator reconstructs the lighting effect of the source site at the viewing site. It processes the lighting control signals received by the presentation functionality in accordance with the conditions unique to the particular presentation environment at the viewing site. The lighting configurator function consists of a light layout translator.

Light layout translator

The light layout translator assigns the lighting devices to lighting control signals so that the lighting effect of the source site is best reconstructed at the viewing site. It may convert and mix the signals to reproduce the lighting effect with the equipment at the viewing site. The assignment may be determined according to the layout of the lighting equipment at the viewing and source sites.

5) Five-sensory configurator

The five-sensory configurator is a function to process all the other information for presentation. It deals with possible future presentation devices for the five senses (sight, hearing, touch, smell, and taste) of the human being. For example, it may control devices to provide a vibro-tactile feeling to the audience. The five-sensory configurator assigns the five-sensory devices to the corresponding information that best stimulates the sensation.

Appendix II

Implementation guidelines for ILE presentation environments using the reference models

(This appendix does not form an integral part of this Recommendation.)

II.1 Implementation guidelines for the proscenium style presentation environment

This clause provides implementation guidelines that describe consideration points for implementing an actual proscenium style presentation environment based on the reference model shown in clause 7.1.

II.1.1 Variety of options

There are a number of options for the reference model of the proscenium style presentation environment, as discussed in clause 7.1. This clause summarizes these options. Figure II.1 shows the various combinations that can be used. The basic model has foreground and background screens. A combination of three screen options, four physical object options, and two sound options creates 24 options, as shown in Table II.1.

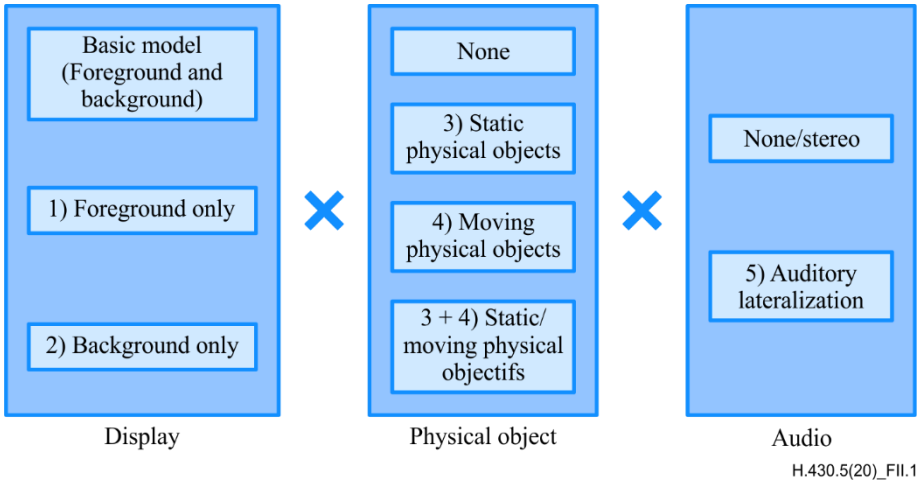


Figure II.1 – Combination of options

Table II.1 – Variety of options on proscenium style reference model

No.	Description	Displays	Static/moving physical objects	Audio
1	Basic model	Both (foreground and background)	None	Stereo
2	Basic model + static	Both	Static	Stereo
3	Basic model + moving	Both	Moving	Stereo
4	Basic model + SM	Both	Static + moving	Stereo
5	Basic model + AL	Both	None	Auditory lateralization
6	Basic model + static + AL	Both	Static	Auditory lateralization
7	Basic model + moving + AL	Both	Moving	Auditory lateralization
8	Basic model + SM + AL	Both	Static + moving	Auditory lateralization
9	P3D	Foreground	None	Stereo

Table II.1 – Variety of options on proscenium style reference model

No.	Description	Displays	Static/moving physical objects	Audio
10	P3D + static	Foreground	Static	Stereo
11	P3D + moving	Foreground	Moving	Stereo
12	P3D + SM	Foreground	Static + moving	Stereo
13	P3D + AL	Foreground	None	Auditory lateralization
14	P3D + static + AL	Foreground	Static	Auditory lateralization
15	P3D + moving + AL	Foreground	Moving	Auditory lateralization
16	P3D + SM + AL	Foreground	Static + moving	Auditory lateralization
17	UHD	Background	None	Stereo
18	UHD + static	Background	Static	Stereo
19	UHD + moving	Background	Moving	Stereo
20	UHD + SM	Background	Static + moving	Stereo
21	UHD + AL	Background	None	Auditory lateralization
22	UHD + static + AL	Background	Static	Auditory lateralization
23	UHD + moving + AL	Background	Moving	Auditory lateralization
24	UHD + SM + AL	Background	Static + moving	Auditory lateralization
Legend: SM: Static and moving physical objects; AL: Auditory lateralization; P3D: Pseudo-3D images; UHD: Ultra High Definition				

Some of the above options may not be applicable, depending on the size and shape of the viewing site. For example, options 17–24 (UHD cases) are not applicable for small venues or small screens.

Figure II.2 illustrates the basic model, screen options 1) and 2), and physical object option 3 + 4) of Figure II.1 from the side view. The basic model in Figure II.2 consists of a foreground transparent screen and a background screen, and the audience watches both screens. Screen option 1) in the figure features one foreground transparent screen located in front of the audience. This option can be used for lectures or speeches, where just one presenter is projected. Screen option 2) in the figure shows a case where only the background screen is used. This option can be used for public or live screening events by using extra-large and ultra-high-definition displays. However, screen option 2) may not be feasible if relatively small displays are used and audiences cannot feel immersed. Physical object option 3+4) in the figure show two physical objects: one in front of the foreground screen and the other located between the foreground and background screens. The objects can be either static or moving physical objects.

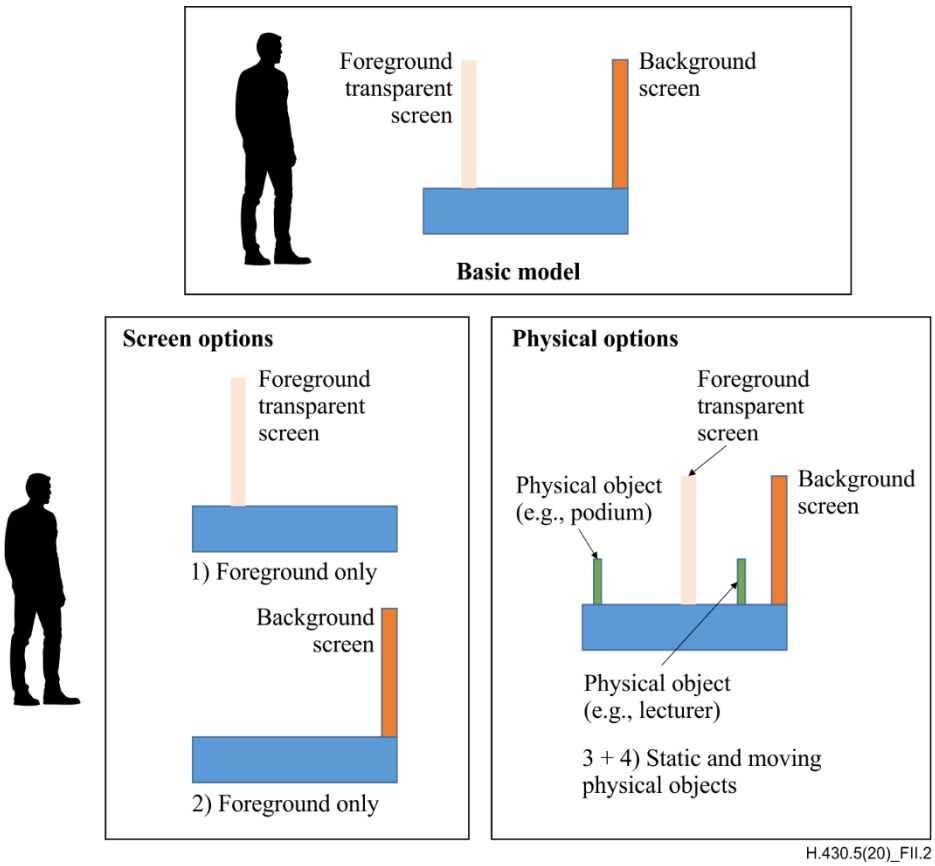


Figure II.2 – Illustrations of screen and physical options

II.1.2 Considerations for projection

The foreground screen can be implemented with any type of transparent display device:

– Virtual image shown using the principle of Pepper's ghost

– Transparent displays (e.g., organic electroluminescent display)

– Transparent diffuser board

– Scrim

The background screen can be implemented with any type of display device:

– Flat wall at one side of the viewing site where images can be projected

– Screen

– One or more flat displays (e.g., multiple 8K displays)

II.1.3 Screen layout example

This clause introduces a screen layout example for a badminton match at the viewing site. Figure II.3 shows an example of the screen layout.

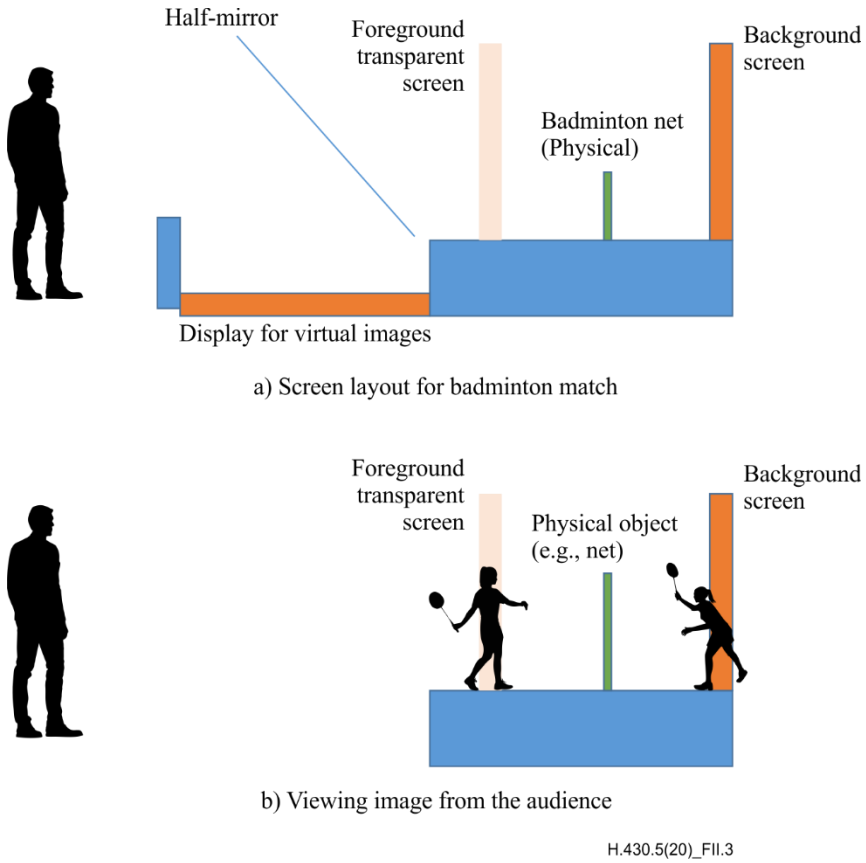


Figure II.3 – Example screen layout for badminton match

An example of the screen layout for a badminton match is shown in Figure II.3, where (a) consists of a foreground and background screens and a badminton net as a static physical object, which is based on the basic model with the static physical object option. The audience can watch the badminton match from one end of the badminton court, as shown in (b).

For the foreground screen, this example uses the Pepper's ghost principle to project pseudo-3D images of one player. Another player is projected on the background screen behind the badminton net, which is located between the foreground and background screens. Information from the event site includes extracted players' images, background images, audio, and their location information to reconstruct the images and sound at the viewing sites. At the viewing sites, one player's image and the background image are displayed on the background screen, and another player's image is displayed on the foreground transparent screen in a pseudo-3D manner using the Pepper's ghost principle.

II.1.4 Considerations for sound/audio

Audio is also very important information for making the audience feel immersed in the presentation environment. The audio information from the event site is transmitted to a multi-input multi-output mixer, and the output signals are distributed to the speakers considering the location of the speakers and the location of the sound source. Such coordination can be performed using existing standards such as [b-ITU-R BS.775-3].

II.1.5 Consideration points for proscenium style presentation environment

There are several consideration points for implementing the proscenium style presentation environment.

II.1.5.1 Consideration points for pseudo-3D images using the Pepper's ghost principle

1) Display layout
   - Tilt angle of the screen should be considered by setting the display at the top or bottom.
   - When applying the moving physical object option, especially with real performers, it is better to place the display at the bottom because this secures a wider space for the performance.
   - It should be confirmed that enough floor space is available when the display is placed at the bottom.
   - It should be considered that the display should not be visible to the audience.
2) Adjusting display position
   - The display position should be adjusted so that the audience can see the pseudo-3D image on the stage by adjusting the display location, image position, stage height, and angle of the half-mirror.
3) Adjusting eye-line
   - The height of the virtual images should be adjusted to the eye-line of the audience.

II.1.5.2 Consideration points for using transparent displays

1) Reflection from other lighting equipment
2) Reflection of audience seats
   - Light and mirror balls
   - Pen lights
   - Camera flashes

II.2 Implementation guidelines for the open style presentation environment

This clause provides implementation guidelines that describe consideration points for implementing an actual open style presentation environment based on the reference model presented in clause 7.2.

II.2.1 Variety of options

There are a number of options for the reference model of the open style presentation environment, as discussed in clause 7.2. This clause explains the options, and Figure II.4 shows their various combinations.

The runway options 1) to 4) in clause 7.2 are classified into two categories by display type: A) a double-sided display, which enables audiences on both sides of a runway to watch images on one screen (Figure 7-3), and B) a single-sided display, which enables audiences to watch images from one side. Runway option 1) can be implemented by A) a double-sided display or several B) single-sided displays, and runway options 2) and 3) can be implemented by B) a single-sided display. By combining A) and B), runway option 4) can be implemented.

Considering other options, display options on the runway stage and front stage, and audio options, 36 combinations of options are possible. In addition, option 9), proscenium style ILE for front stage, has 24 options (as described in clause II.1.1).

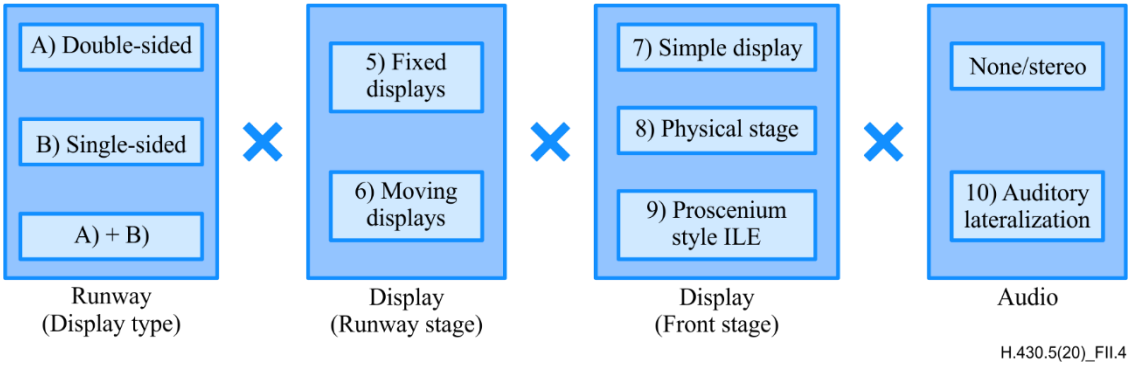


Figure II.4 – Combinations of major options

Since the open style can feature one or more runways in the middle of the hall, audiences can see the back view of objects from the back side screen by using A) the double-sided option. Physical options for the runway stage are the same as those in the proscenium style presentation environment.

II.2.2 Considerations for projection

When the proscenium style presentation is used for the front stage, the consideration points are the same as those of the proscenium style presentation environment described in clause II.1.2.

For runway stages, the transparent screen can be implemented by any of the devices for the foreground screen of the proscenium style presentation environment:

– Virtual image shown using the principle of Pepper's ghost
– Transparent displays (e.g., organic electroluminescent display)
– Transparent diffuser board
– Scrim

The background screen of the proscenium style presentation environment is not typically used for runway stages.

II.2.3 Screen layout example

This clause introduces a screen layout example for runway stages at the viewing site. Figure II.5 shows two examples of the screen layout.

Figure II.5(a) shows an example of a screen layout for a double-sided display, which consists of a transparent screen. Audiences on both sides can watch different images: for example, audiences on the left side see the front view of the displayed performer and audiences on the right side see his/her back view simultaneously. As an implementation, the double-sided display can be achieved by combining two transparent displays back to back.

Figure II.5(b) shows an example of a screen layout for a single-sided display, which consists of a transparent screen. Audiences can watch images displayed on the screen from a single side. A traditional opaque screen can also be used for the single-sided display, but the transparent screen is usually preferred because it enhances immersiveness by merging the displayed images into the viewing site.

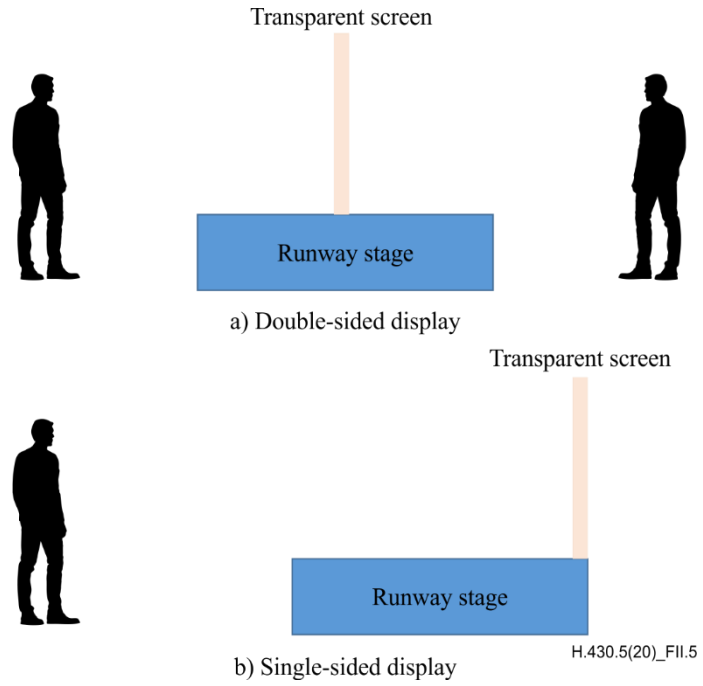


Figure II.5 – Two screen layout examples for runway stage

In the open style presentation environment, there are moving display options on runway stages. Figure II.6 shows two examples of the fixed display option: (a) an example on runway stages and (b) an example hung from the ceiling over the runway stage. These types can be selected on the basis of conditions at the viewing sites such as height of the ceiling, capacity of the hall, and size of the displays.

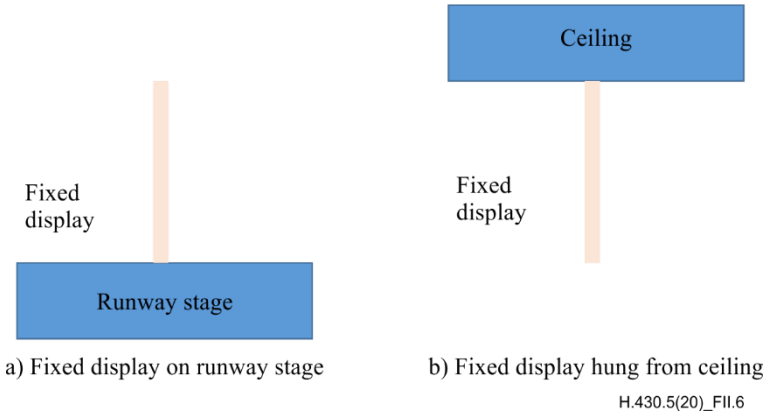


Figure II.6 – Screen layout examples for fixed display option

The moving display option is a characteristic option of the open style presentation environment. An example is shown in Figure II.7.

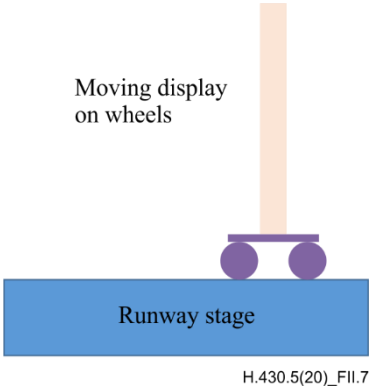


Figure II.7 – Screen layout example for moving display option

Displays that are mounted on a dolly with wheels can be moved along the runway stages. Either a double-sided or single-sided display can be selected. With the double-sided option, the audience can feel a stronger sense of immersiveness when performers such as musicians are displayed.

II.2.4 Consideration for sound/audio

The consideration points for sound/audio are the same as those for the proscenium style presentation environment.

II.2.5 Consideration points for the open style presentation environment

There are several consideration points for implementing the open style presentation environment. Most of the consideration points for the proscenium style presentation environment are applicable to the open style.

1) Display layout
   - Projection images might require keystone correction when a transparent screen is used, as projectors are typically located on the floor or the ceiling. It should be noted that the resolution of the projected images might be lower than the specs of the projector.
   - When projectors are located on the floor, they should be carefully positioned so as not to be occluded by the audience.

2) Adjusting display images
   - When using the double-sided display option, it might be required to adjust the position and size of the displayed images, and to synchronize the front and back views.
   - When using screen stitching of several displays, synchronization, position adjustment, and size adjustment of the displayed images are required.

3) Transparent screens on runway stages
   - Consideration should be given to the transmission image. Audiences on the back side might see the front images through the transparent screen.
   - Care should be exercised not to let the direct light from the projector get into the audiences' eyes on the other side of the transparent screen.
   - Since the image is projected not only on the screen but also through the screen, care should be taken that this undesirable projection not be visible to the audience.
   - Since the scrim and/or screens may sway due to factors such as air conditioning, the audience's footsteps, and the vibrations of the performers on the stage, means to suppress vibration and shaking may be needed.

II.3 Implementation guidelines for arena style presentation environment

This clause provides implementation guidelines that describe consideration points for implementing an actual arena style presentation environment based on the reference model shown in clause 7.3.

II.3.1 Variety of options

There are several options for the reference model of an arena style presentation environment, similar to the proscenium style. Since the arena style has four screens, there are no background screens to enable audiences to see the back side view of objects from the back. Therefore, the display option is limited to "foreground only" in Figure II.1. Other options, both physical options and sound options, are the same as those for the proscenium style presentation environment.

As an example of a physical option, Figure II.8 shows a collaboration with a real performer as a moving physical objects option for the arena style presentation environment. In this case, the viewing site must be large enough and with a high enough ceiling to accommodate a large four-sided display structure in which real performers can interact with the reflected images.

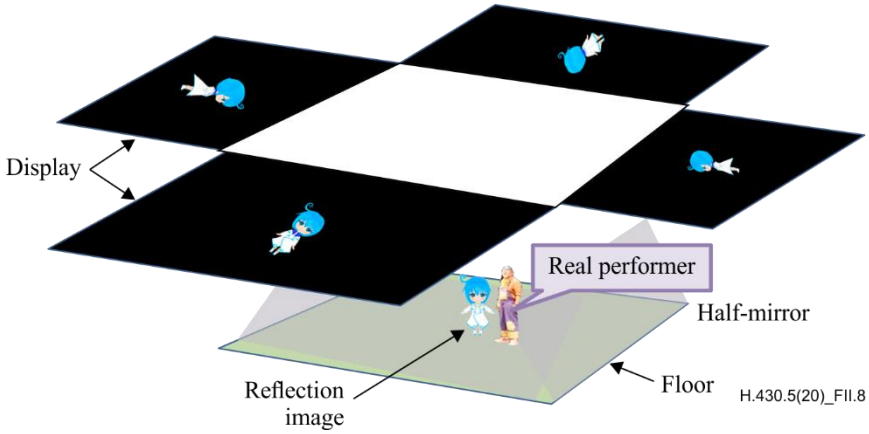


Figure II.8 – Physical object option of arena style

II.3.2 Considerations for projection

The foreground screen can be implemented with any type of transparent display device:

– Virtual image shown using the principle of Pepper's ghost

– Transparent displays (e.g., organic electroluminescent display)

– Transparent diffuser board

– Scrim

The background screen is typically not used.

II.3.3 Screen layout example

This clause introduces screen layout examples for the arena style presentation environment. Figure II.9 shows examples of the screen layout.

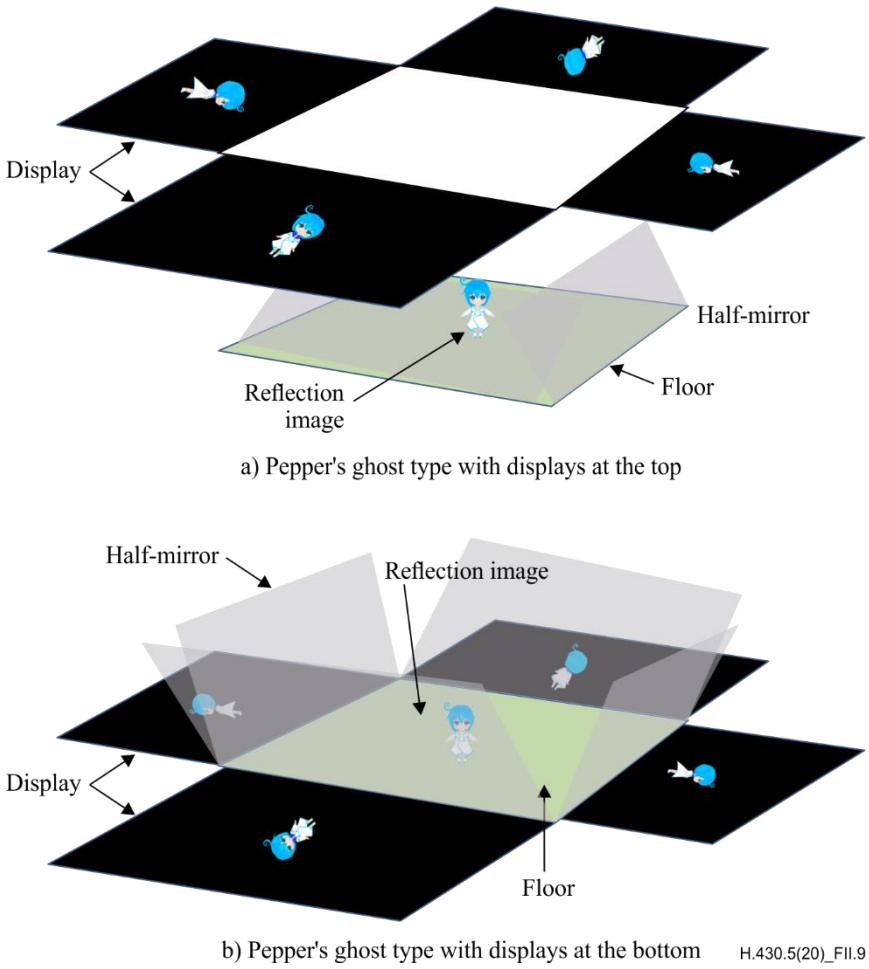


Figure II.9 – Examples of arena style screen layout

Each example of the screen layout for the Pepper's ghost type shown in Figure II.9 consists of four displays and four half-mirrors. The displays can be placed either at the top or bottom, with the half-mirrors tilted inwards and outwards as shown in Figure II.9(a) and II.9(b), respectively. Audiences can watch the scene from all directions. Information from the event site includes extracted performer's images from four directions, audio, and the location information to reconstruct the images and sound at the viewing sites. At the viewing sites, the images are shown on the four displays and reflected on the half-mirrors in a pseudo-3D manner using the Pepper's ghost principle.

II.3.4 Considerations for sound/audio

Consideration points for sound/audio are the same as those for the proscenium style presentation environment.

II.3.5 Consideration points for arena style presentation environment

There are several consideration points for pseudo-3D images using the Pepper's ghost principle, as one way of implementing the arena style presentation environment.

1) Display layout

- Tilt angle of half-mirrors should be considered in light of whether the displays are placed at the top or bottom.
- Care should be taken that the display not be visible to the audience.

2) Adjusting display position

- Display position should be adjusted so that the audience can see the pseudo-3D image on the stage by adjusting display location, image position, stage height, and angle of the half-mirror.
- Positions of all displays should be adjusted so that the reflected images are crossed orthogonally at the centre of the stage. This means that the reflected images of the front and back views and the left and right views are projected back to back at the same position.

3) Adjusting eye line

- The height of the aerial images should be adjusted to the eye line of the audience.

4) Brightness of viewing site

- As there is an audience all around the stage in the arena style presentation environment, the audience watches the projected images not only from the front side but also the back, left, and right sides. Usually the circumference of the audience is brightened up, but the background of projected images should be dark to take into account the principle of Pepper's ghost. In order to achieve a good balance between these, the brightness at the viewing site should be adjusted so that the audience is not bothered by the brightness of the background of reflected images and the other audience members on the opposite side.

Bibliography

[b-ITU-R BS.775-3] Recommendation ITU-R BS.775-3 (2012), *Multichannel stereophonic sound system with and without accompanying picture*.

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

Printed in Switzerland
Geneva, 2020