International Telecommunication Union

ITU-T Technical Paper

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(30 April 2021)

HSTP.ACC-UC
Use cases for inclusive media access services

Summary

This Technical Paper describes use cases for multimedia accessible system, and in particular an experiment of Internet protocol television (IPTV) services with accessibility functions based on Recommendation ITU-T H.702 about accessibility profiles for IPTV systems.

Note

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

Keywords

Accessibility, accessibility profiles, IPTV, multimedia, use cases.

Change Log

This document contains Version 1 of the ITU-T Technical Paper HSTP.ACC-UC "*Use cases for inclusive media access services*" approved at the ITU-T Study Group 16 meeting held online, 19-30 April 2021.

Editor:	Hideki Yamamoto	Tel:	+81(0)48 420 7012
	Oki Electric Industry Co., Ltd.	Email:	yamamoto436@oki.com
	Japan		

Table of Contents

Technical Paper ITU-T HSTP.ACC-UC

Use cases for inclusive media access services

1   Scope

This Technical Paper describes use cases for multimedia accessible systems and their use cases. It also describes an experiment of Internet protocol television (IPTV) services with accessibility functions based on [ITU-T H.702] on accessibility profiles for IPTV systems.

2   References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Technical Paper. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Technical Paper 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 Technical Paper does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T F.791]   Recommendation ITU-T F.791 (2018), *Accessibility terms and definitions*.

[ITU-T H.702]   Recommendation ITU-T H.702 (2020), *Accessibility profiles for IPTV systems*.

[ITU-T H.721]   Recommendation ITU-T H.721 (2015), *IPTV terminal devices: Basic model*.

[ITU-T H.760]   Recommendation ITU-T H.760 (2009), *Overview of multimedia application frameworks for IPTV services*.

[ITU-T H.761]   Recommendation ITU-T H.761 (2014), *Nested context language (NCL) and Ginga-NCL*.

[ITU-T H.762]   Recommendation ITU-T H.762 (2011), *Lightweight interactive multimedia environment (LIME) for IPTV services*.

ITU-T H.763.1   Recommendation ITU-T H.763.1 (2010), *Cascading style sheets for IPTV services*.

[ITU-T H.764]   Recommendation ITU-T H.764 (2019), *IPTV services enhanced script language*.

[ITU-T H.765]   Recommendation ITU-T H.765 (2015), *Packaged IPTV application (widget) service*.

[ITU-T Y.101]   Recommendation ITU-T Y.101 (2000), *Global Information Infrastructure terminology: Terms and definitions*.

[ITU-T Y.1901]   Recommendation ITU-T Y.1901 (2008), *Requirements for the support of IPTV services*.

[ITU-T Y.1910]   Recommendation ITU-T Y.1910 (2008), *IPTV functional architecture*.

3   Definitions

3.1   Terms defined elsewhere

None.

3.2 Terms defined in this Technical Paper

None.

4 Abbreviations and acronyms

This Technical Paper uses the following abbreviations and acronyms:

IPTV   Internet protocol television

ISDB-T   Integrated Services Digital Broadcasting-Terrestrial

ITA   IPTV Terminal device with Accessibility functions

PWD   Persons with Disabilities

STB   Set-top box

TV   Television

5 Introduction

The United Nations Convention on the Rights of Persons with Disabilities (PWD) [b-UNHR] was adopted by the United Nations General Assembly on 13 December 2006, and came into effect in 2007. This convention requires the ratifying countries to ensure that persons with disabilities enjoy full equality under the law. It also explicitly requires them to make provisions for information accessibility of the PWD in some ways.

To disseminate the use of accessible systems based on global standards, it is important to share use cases based on standards such as the ITU-T standards listed in clause 2.

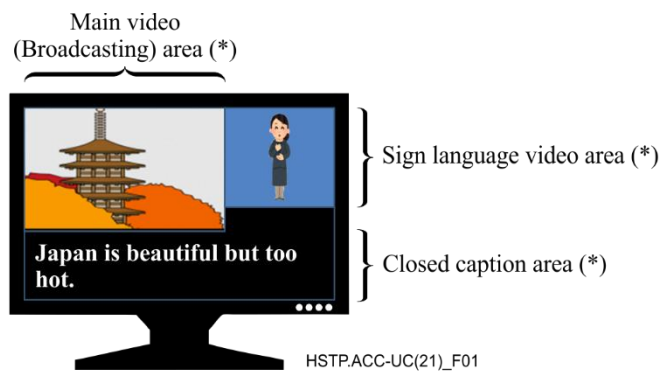This Technical Paper describes use cases of accessibility enhanced system based on ITU-T standards.

6 Use cases of multimedia accessible systems based on global standards

6.1 Case 1: Broadcast TV and IPTV system based on ITU-T H.702

Broadcast television (TV) terminal is a one-way device to get information. From the viewpoint PWDs, sufficient accessibility information is not attached to the content. In the case of IPTV, with interactive multimedia application frameworks, it is possible to easily add a variety of accessibility information on the video. [ITU-T H.702] is a global standard for "Accessibility profiles for IPTV systems." It defines three profiles for accessibility features in IPTV systems, with increasing levels of support.

In this use case, video is delivered by broadcasting and accessibility information such as captions, sign language streams and audio description are sent separately from video contents to IPTV terminal devices. It is not necessary that accessibility information is delivered by broadcasting companies. It is possible that the accessibility service provider, a kind of IPTV service provider, provides them based on the profiles defined in [ITU-T H.702]. The terminal device that is ITU-T H.702 compatible can receive both broadcasting video and accessible information and display them as shown in Figure 1.

This use case is experimented and the results in detail are described in Appendix I. In summary, the delay between video and accessibility information on the terminal device was measured by hand, yielding a delay of a few seconds, approximately. In detail, the delay of sign language was four seconds. It included the translation by sign language translators, video delivery from the studio and set-top box (STB) and video decoding. The participants of this experiment commented that the delay was acceptable for understanding the video.

Main video
(Broadcasting) area (*)

Sign language video area (*)

Closed caption area (*)

Japan is beautiful but too hot.

HSTP.ACC-UC(21)_F01

(*) Size and position of both sign language and closed caption can be changed by remote controller

Figure 1 – Layout of contents on the display

Appendix I

An experiment of IPTV services with accessibility functions based on ITU-T H.702

I.1   Background

The United Nations Convention on the Rights of Persons with Disabilities (PWD) was adopted by the United Nations General Assembly on 13 December 2006, and came into effect in 2007. This convention requires the ratifying countries to ensure that persons with disabilities enjoy full equality under the law. It also explicitly requires them to make provisions for information accessibility of the PWD in some ways.

Television (TV) is a one-way device to get information. From the viewpoint of PWDs, sufficient accessibility information is not attached. In the case of Internet protocol television (IPTV), with interactive multimedia application frameworks, it is possible to easily add a variety of accessibility information on the video. [ITU-T H.702] is a global standard for "Accessibility profiles for IPTV systems." It defines three profiles for accessibility features in IPTV systems, with increasing levels of support. While the basic profile provides an entry-level support of accessibility, the main profile provides the widest range of features. The enhanced profile provides the middle level support between the basic profile and the main profile. Accessibility information is information such as captions, sign language streams and audio description that are sent separately from video contents to IPTV terminal devices. By defining the above profiles, persons with disabilities can choose more easily the terminal devices that have the functions they need. After publishing [ITU-T H.702], [b-TTC JT-H702] was published as a Japanese Standard by TTC. [b-TTC JT-H702] is so called down-stream standard because it is based on [ITU-T H.702]. The set-top boxes (STBs) and servers that support [ITU-T H.702] were developed but had not been tested by the potential users.

This Technical Paper reports the result of the experiment of IPTV systems that supports [ITU-T H.702] with PWDs. The experiment was conducted at Kyoto in Japan, on 30 July 2017. Participants of the experiments were potential users of this system, members of organizations of persons with disabilities, a broadcasting operator, an IPTV service operator, and experts of accessibility standards. They watched the TV program with sign language video and caption and used the functions to enlarge/shrink the size of sign language video and caption and so on. After watching the test program, they discussed it.

In the discussion, the usability of the function that displays sign language video on the screen with the broadcasting program was confirmed by participants.

I.2   Introduction to ITU-T H.702

[ITU-T H.702] organizes the functions that satisfy the following four general requirements and divides the said functions into four components to define the profiles.

a)   IPTV terminal device with accessibility functions (ITA) is recommended to support synchronization of streams in accessibility medium (e.g., caption, sign-language and audio description) with the main video content.

b)   ITA is recommended to support the optionality of accessibility information (e.g., to turn on/off displaying).

c)   ITA is recommended to support changes in media-positioning (e.g., to change the position of caption and sign language, or to change the spatial location of audio description).

d)   ITA is recommended to support changes in media properties (e.g., audio-volume, text-scrolling, font-size, font-colour, video-frame rate, etc.).

A requirement that recommends ITA to support personalized profiles according to user preferences, to allow the user to choose among the functionalities of the supported profile is for future study.

I.2.1   Architecture and capability of ITU-T H.702

Figure I.1 shows the architecture of how to get the accessibility medium within a derivative of IPTV architecture (defined in [ITU-T Y.1910]). Caption, sign language and audio description are delivered by accessibility medium server functions. The end-user can display the accessibility medium on terminal devices by using accessibility medium terminal functions.
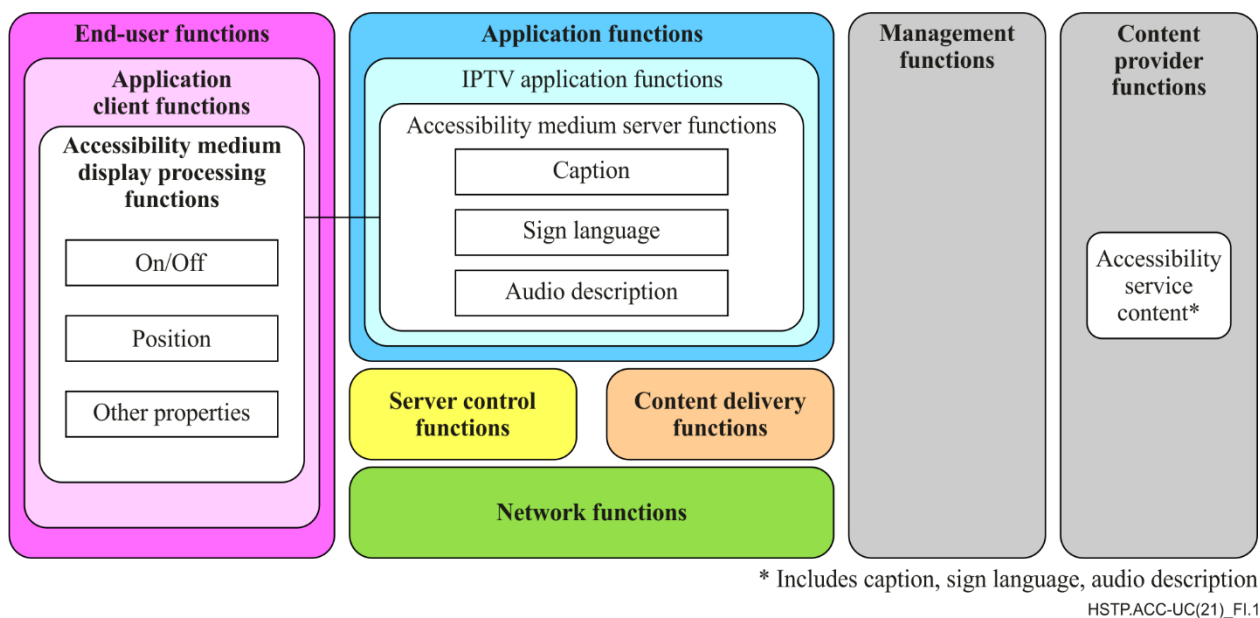

* Includes caption, sign language, audio description

HSTP.ACC-UC(21)_FI.1

Figure I.1 – Functional blocks of accessibility service

It is expected that application client functions be implemented by multimedia application framework such as [ITU-T H.760], [ITU-T H.761] Ginga-NCL, [ITU-T H.762] LIME, [ITU-T H.763.1], [ITU-T H.764] and [ITU-T H.765].

Table I.1 shows capabilities of caption, sign language and audio description that are defined in [ITU-T H.702].

Table I.1 – Capabilities of caption, sign language and audio description

Accessibility Medium	Capabilities
Caption	Turn on/off overlaid caption
	Change the directions of displaying text between horizontal and vertical
	Change the transition effects of the caption text between cut and scroll
	Select from multiple captions
	Change font size of the caption text
	Change font style of the caption text
	Change font colour of the caption text
	Change caption position from overlaid (on-screen) or off-screen
	Change the background colour of caption box
	Change the size of caption box

**Table I.1 – Capabilities of caption, sign language and audio description**

Accessibility Medium	Capabilities
	Synchronize caption with video
	Synchronize caption with the main video during the playback mode including slow motion
	Generate automatically multiple captions with speech recognition
	Display caption to different display devices
	Change display speed of the caption text
	Hold the language setting of a caption when multiple captions are supported (Function of holding the language setting. When one changes channels, one can get the same language caption as in the previous channel.)
Sign language	Turn on/off overlaid sign language
	Select from multiple sign languages
	Change video size of sign language
	Change video position of sign language
	Change the background colour of sign language video
	Synchronize sign language video with the main video
	Synchronize sign language video with the main video during the playback mode including slow motion
	Generate automatically synthesized sign language interpretation
	Avoid covering important information of original main video with sign language video
	Hold the language setting of sign language when multiple sign languages are supported
	Turn on/off audio description
Audio description	Synchronize audio description with the main video
	Adjust volume of audio description
	Adjust sound quality of audio description
	Synchronize audio description with the main video during the Playback mode including slow motion
	Avoid interfering original main audio with audio description
	Select from multiple audio descriptions
	Hold the language setting of audio description when multiple audio descriptions are supported
	Read letters and description of button on the screen

These functions are recommended to be implemented not only separately but also in combination (e.g., captions and sign language).

I.3 Experiment system architecture

Figure I.2 shows the architecture of an experiment system. The terrestrial broadcasting program used in this experiment was a real program. It was broadcasted from the broadcasting studio in Kyoto. Research participants in the service area watched the program by set-top box (STB) that support [ITU-T H.702] (hereafter H702 STB). The STB received not only broadcasting signal through

antenna but also IP signal through broadband network. A studio of TV for PWD also received the broadcasting signal as other research participants and translated it to sign language. The video of sign language was encoded and transmitted to a sign language server (IPTV server) in real time. The sign language server delivered the sign language video to a managed network, one of the broadband network in which the quality of service is managed by a service operator. Caption of the program was generated in the studio of TV for PWD, too. It was delivered from caption server through the Internet.
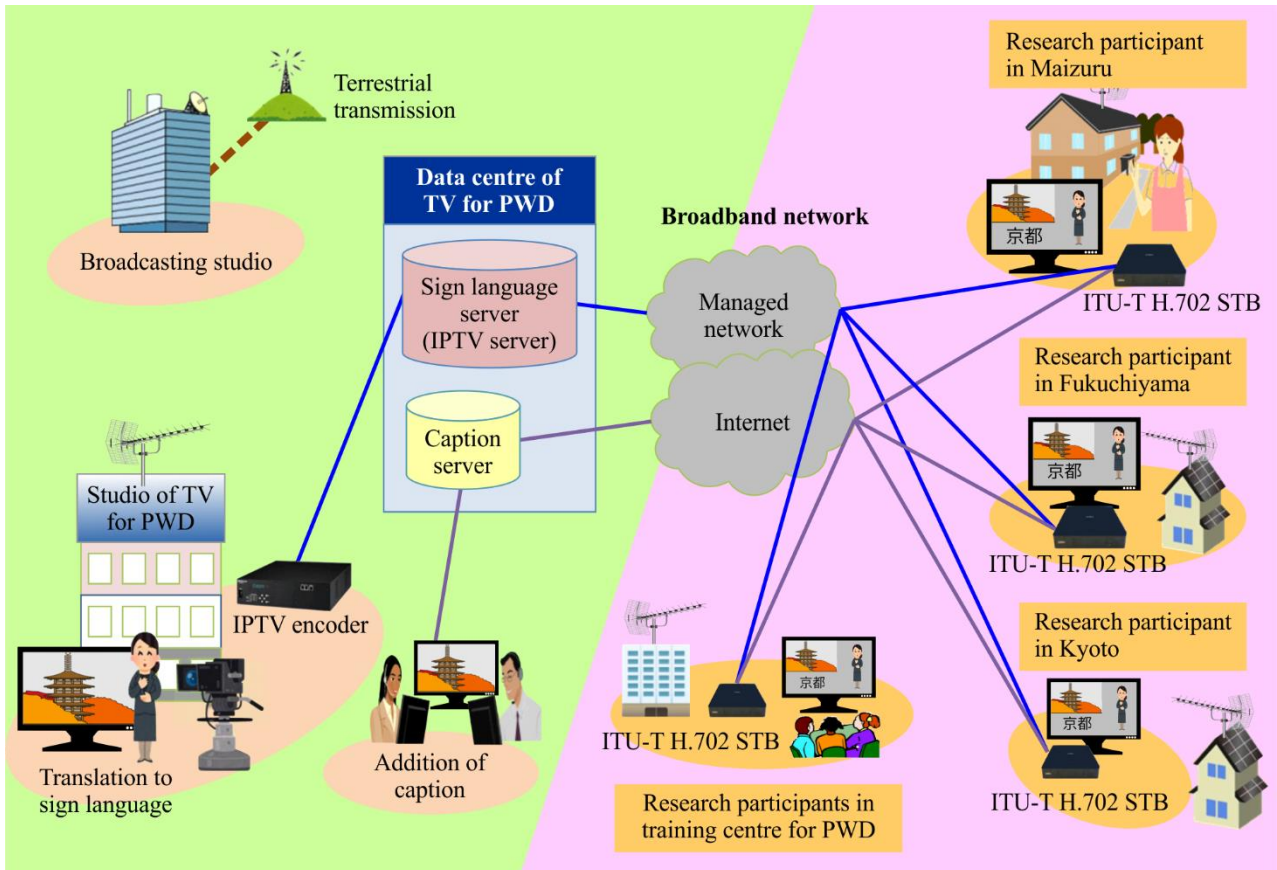
HSTP.ACC-UC(21)_FI.2

Figure I.2 – Experiment system architecture

Figure I.3 shows the layout of contents on the display in this experiment. It is an output of ITU-T H.702 STB. The broadcasting area shows the original broadcasted programme used in its original state. Sign language video area shows the translation of the program into sign language. In this architecture, sign language is generated by translators while they watch the program. Then, it is encoded and delivered through broadband network. It means the delay of sign language happens on the display. Caption area shows the caption of this program from caption server. As [ITU-T H.702] defines, the layout of Figure I.3 can be changed through a remote controller. For example, caption colour can be changed, and the size of sign language area can be enlarged.
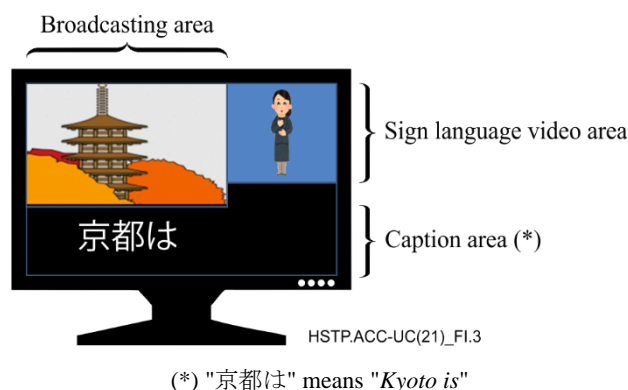
(*) "京都は" means "*Kyoto is*"

Figure I.3 – Layout of contents on the display

I.4 Experiment detail

The detail of the experiment is as follows:

– Organizer: Approved Specified Non-profit Corporation Japanese Organization of Broadcast and Communications for People with Disability and Kyoto Broadcasting System Company Limited

– Cooperation: Japanese Federation of the Deaf, IPTV Accessibility Consortium,

  • All Japan Association of Hard of Hearing and Late-Deafened People,

  • Kyoto-fu association of hearing-impaired people,

  • Kyoto-fu association of hard of hearing people,

  • Kyoto city association of hard of hearing and late-deafened people,

  • Kyoto welfare association of hearing-language-impaired people.

– Date: 19:00 – 19:55 (JST), 30 July 2017

– Number of sites for research participants: 5

– Number of participants: approximately 30

– Broadcasting program: Okashina Kyoto (Refined Kyoto cake) by Kyoto Broadcasting System Company Limited

– Accessibility information: Federation of broadcasting for person with disabilities and ASTEM Co., Ltd.

– Platform specifications:

  • Broadcasting system: ISDB-T

  • Sign language server: [ITU-T H.721] (OKI MediaServer by Oki Electric Industry Co., Ltd.)

  • Caption server: Web server (ASTEM)

  • STB: IPTV and RF hybrid STB ([ITU-T H.702], [ITU-T H.721], [ITU-T H.762], AI-dragon 4 by ASTEM., Co. Ltd.)

  • Bit rate of sign language video: 3 M/bit/s

  • Resolution of sign language video: 1280×720 / 1920×1080

- Delay between video and accessibility information:

    The delay between video and accessibility information on the terminal device was measured by hand, yielding a delay of a few seconds, approximately. In detail, the delay of sign language was four seconds. It included the translation by sign language translators, video delivery from the studio and STB and video decoding. The participants of this experiment commented that the delay was acceptable for understanding video.

Bibliography

[b-UNHR]   United Nations Convention on the Rights of Persons with Disabilities
                  https://www.ohchr.org/en/hrbodies/crpd/pages/conventionrightspersonswithdisabilities.aspx
                  (visited 2021-05-28)

[b-TTC JT-H702]   The Telecommunication and Technology Committee JT-H702 version 1 (10
                  November 2016), *Accessibility profiles for IPTV systems* (in Japanese).
                  https://www.ttc.or.jp/application/files/1815/5426/8315/JT-H702v1.pdf (visited 2021-05-28)

_____