ITUPublications

International Telecommunication Union
Radiocommunication Sector

Report ITU-R BT.2448-1

(09/2025)

BT Series: Broadcasting service (television)

Technical realisation of signing in digital television

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from https://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Reports

(Also available online at https://www.itu.int/publ/R-REP/en)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	Broadcasting service (sound)
BT	**Broadcasting service (television)**
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radio-wave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management
TF	Time signals and frequency standards emissions

***Note**: This ITU-R Report was approved in English by the Study Group under the procedure detailed in Resolution ITU-R 1.*

Electronic Publication
Geneva, 2025

REPORT ITU-R BT.2448-1

Technical realisation of signing in digital television

(2019-2025)

TABLE OF CONTENTS

1 Introduction

In 2006, the United Nations released the Convention on the Rights of Persons with Disabilities (CRPD)[1]. The Convention entered into force on 3 May 2008. The CRPD especially recognizes and promotes the use of sign languages. The scope of this Report is to describe methods of making sign language available to the end users.

Many broadcasters already provide signing for people that are hard of hearing and who depend on sign language for their communications. The sign-language interpreter is usually captured by a separate camera.

The signer video can be inserted at the studio (open signing) or can be streamed separately within the TV multiplex, for example as additional programme elementary stream, or via the Internet. The separate stream enables closed signing, i.e. the consumers can switch the signer video on or off. The signers can be real persons or computer-generated sign language interpreters (avatars).

In the case where a TV set or set-top box is equipped with a second video player, the streamed signer video can be decoded in the device and, after time synchronisation, then superimposed on the television broadcast programme. Otherwise, the streamed signer video can be displayed (after time synchronization) on a second (i.e. separate) screen, for example on a tablet PC or computer monitor placed next to the main TV screen or by means of video glasses to create a virtual image of the signer video.

In the case where avatars are used for sign language interpreters, avatars can be rendered on the broadcaster's side or on the receiver side (or companion device) based on the sign language information transmitted for controlling the avatars.

2 Open signing versus closed signing

Analogous to open and closed captions, the signer videos can be inserted permanently in the TV programme (open signer) or can be made available in a separate video stream, either within the TV multiplex, for example as an additional programme elementary stream, or via the Internet. The separate stream enables closed signing, i.e. the consumers can switch the signer video on or off.

Open signers are visible to anybody and cannot be switched off by the end-user. When closed, the signer videos are made available on demand, i.e. they can be switched on and off by the end-user.

People who do not depend on the presence of a sign-language interpreter often feel disturbed by the permanently superimposed signer video. This situation is comparable to open subtitles/captions that may bother viewers who do not need or like them. Consequently, broadcasters everywhere in the world are currently attempting to offer the signer video in closed form.

3 Methods of providing closed signing

3.1 Delivery of video streams of signers (real persons or avatars)

Principally, there are two methods of providing closed signing using video streams:

a) The broadcaster provides two video signals, one with and one without the signer: The consumers can select either programme. The programme with the inserted signer is created in the studio (studio mix) and is either included in a digital TV programme multiplex or made available as IP stream via the Internet (see Fig. 1).

---

[1] https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html

b) The broadcaster provides the signer video as additional stream: The signer video is either included in the broadcast multiplex (for example as an additional packetized elementary stream (PES)), or is made available in interactive form via an additional IP connection. In the receiver, the signer video can be superimposed onto the main TV signal (receiver mix) or, alternatively, be displayed on a second screen conveniently placed next to the main screen.

In case a), the viewers have neither the possibility to change position and size of the signer nor to modify the tilt angle of the main TV video signal but may select from several versions (if offered).

In case b), viewers can modify the position and the size of the signer. The inclination of the main TV programme may be fixed by the receiver manufacturer (could be zero) or may be made adjustable by the end-user (from zero to a given maximum value).[2] Figures 2 and 3 provide examples of how the sign-language interpreter may be represented on the TV screen.

FIGURE 1

Example of simulcast approach: German ARD evening news of 3 March 2017

Report BT.2448-01

Left: 1st German Television (no signer) Right: Same news in TV programme "phoenix" (signer)

NOTE 1 – Apart from the 20:00 hour's news, the content of the two programmes is different.

NOTE 2 – The signer is relatively large, has dark clothes and is placed to the lower right of the TV image which is tilted (rotation of the TV image is preferred for news, however, not for sports or documentaries).

3.1.1 Selection of an alternative video signal

In its very basic form, this solution works with any existing digital system. In this case, the signer is inserted in the ongoing TV programme at the studio and then either included as additional service in the broadcast bouquet or offered as Internet stream. Several versions can be created of this alternative video programme (signer at different positions or size, different size and inclination of the ongoing TV programme). This may be realized for Internet streams to allow the users a selection according to their needs and tastes (see Fig. 4).

---

[2] Customization of the signer is a desired feature as a relatively large image of the signer is preferred but sometimes the signer image needs to be moved or made smaller in case it is overlapping with important parts of the TV image, for example with a table of sports results. If the signer video is inserted in the studio and thus has a fixed position and size, tests carried out with sign-language users, undertaken within the EU project Hbb4All, have revealed that there is a significant preference for a rather large signer being placed into the lower right of the TV image. The signer should partly overlap with the TV programme which is inclined for news but not for sports or documentaries (where the image of highest importance). For more details on these studies see, for example, the findings of the EU project Hbb4All [3], [4], [5].

FIGURE 2

Example of relatively large image of the video signer and relatively small and strongly tilted image of the main video (one possible option, often applied for news)

Report BT.2448-02

(1) Alternative video in the broadcast multiplex

The TV programme with the signer can be included in the same broadcast mux as the main TV programme or in a different one also available to the TV receiver (TV set or set-top box). However, if included in a different mux (compared to the main TV programme), all other broadcast elements such as sound or videotext need to be retransmitted, too. So, it is advisable to include the signer service in the same multiplex as the main TV programme it relates to.

Of course, the price to pay is the data capacity occupied by this additional video stream (typically 2-3 Mbit/s or more, depending on the desired picture quality (SD or HD) and the applied video coding system (H.264 or H.265).

It is usually needed to signal to the end-users the availability of an alternative video programme with inserted signer. This can be done by an open text message or an open icon but more elegantly by an icon contained in the launching application of an IBB system such as HbbTV or Hybridcast (see Recommendation ITU-R BT.2075 [1] and Report ITU-R BT.2267 [2]). For example, HbbTV can trigger an autostart application if signalled in the AIT (application information table of the MPEG-2 TS). Once an autostart application is triggered, the user is informed about all available accessibility services. Eventually, remote controls may have a specific button for closed signing (like today for subtitles).

In case the TV programme with the signer is included in a broadcast multiplex and is selected via the launching function of HbbTV (either by clicking on the corresponding icon or by clicking the indicated colour button), the TV programme with the inserted signer can optionally be broadcast in a "hidden" mode which means that it is only available in connection with the main programme and cannot thus, for example, be found in a conventional programme search.[3] Such course of action is useful in case the additional video with the signer is only available part-time. For example, a broadcast mux may regularly contain five HDTV programmes. For one of these programmes, a copy with the signer is made available from time to time in the same multiplex which then carries six programmes.

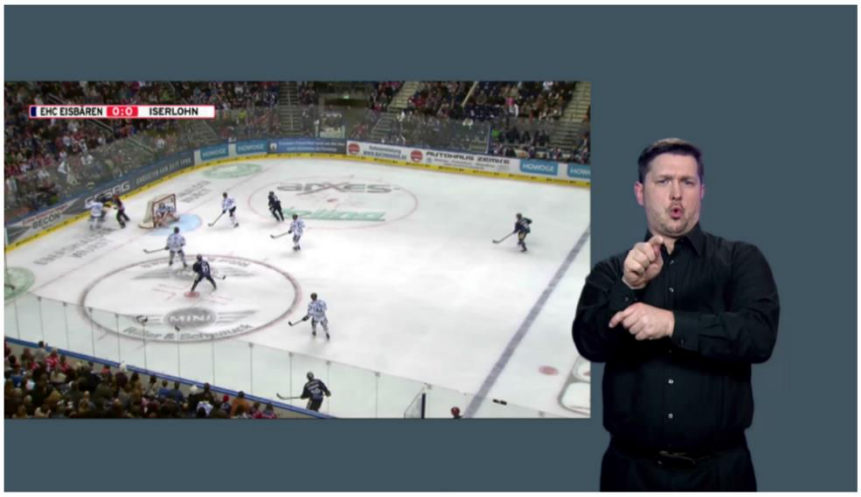
The data rate needed for this sixth TV programme is taken away from the five programmes normally contained. This may result in a temporarily somewhat lower quality for all programmes but if statistical multiplex is applied, this should not make too much of a difference.

---

[3] In DVB, a programme is "hidden" if it is not listed in the Service Description Table (SDT).

FIGURE 3

Example for a medium-sized image of the video signer and relatively large, non-tilted image of the main video with slight overlapping of both images (another possible option)



Report BT.2448-03

(2) Alternative video as Internet stream

Often, broadcasters provide the additional video stream with the signer via the Internet. Interactive TV sets or set-top boxes that are linked to a broadband IP connection (so called "Connected TV" or "Smart TV" sets) can then display this Internet stream. The advantage is that no broadcast data capacity is needed for the video programme with the inserted signer, however, the users need to have a TV set that is connectable and connected to the broadband Internet. Furthermore, the broadcaster has to pay the broadband service provider for the ingestion of the signal. There may also be additional cost for the consumers in case the data rate of the video programme with the signer exceeds the data limit of a consumer's contract with the broadband network provider (which could be especially relevant in case of mobile broadband connections). In the same way as above, use can be made of IBB systems to indicate and launch the IP video that includes the signer.

FIGURE 4

Selection of user preferences: Graphical user interface in a field test carried out by the German public service broadcaster "rbb"

Report BT.2448-04

Note to Fig. 4: Actually marked is the possible spatial arrangement "not overlapping, main TV inclined. Other items for selection are the size and the position of the sign-language interpreter.

3.1.2 Reception of the signer video only (parallel to the TV broadcasting signal)

In this case, the signer video is transmitted separate to the related main TV programme. At the consumer site, the signer video is either overlaid with the main TV programme or is displayed on a second screen. The signal is either transmitted within the broadcast mux or is made available as Internet stream via a broadband network.

The signer video is usually captured with a camera in the studio and then video compressed. The signal may be accompanied by an alpha-channel to allow alpha blending in the receiver. The alpha-channel allows to mix the image of the signer by his or her contours. Without such an alpha channel, the signer video would appear in a separate window.

The advantages of a separate signer video are twofold:

– The data rate of this signer video can be kept substantially lower than the one needed for a composite video that includes both the main TV programme and the signer video.

– The presentation of the signer video can be accomplished according to the preferences of the users (position and size of the signer video).

However, there are technical challenges which are not (yet) mastered by most of the existing TV devices. The receiver mix is only possible in case the TV set or set-top box allows for the simultaneous decoding of two video signals and for the synchronization of these two video signals prior to their display. Synchronization is no problem if the signer video is inserted in the broadcast mux. The second video player, however, is optional in the specifications of HbbTV 2.0. Furthermore, in case of Internet stream, the receiver needs to cope with a delay of some seconds (could be up to 10 seconds) for the signer video. The specification of HbbTV 2.0 provides optionally for a minimum of some 30 MB of cache for that purpose.[4] If we assume a data-rate of 8 Mbit/s (i.e. 1 MB/s) for the main TV video, a delay of up to 30 seconds could thus be compensated.

Furthermore, the receiver would have to provide an engine for the overlaying of the signer video (picture-in-picture function) within a separate window or by alpha blending. For the latter case, the receiver needs to understand the alpha signal and the broadcaster needs to provide the alpha-channel together with the signer video stream. In theory, the MPEG-TS is capable of carrying an alpha-signal, however, up to now, broadcasters have no experience with the transmission of the alpha channel. An alternative to the alpha channel could be to agree on a standardized colour (e.g. a specific hue of green) for the chroma key compositing (“blue-box” procedure) that is used by the broadcaster in capturing the signer video and is known to the receiver manufacturers.

It may be desirable that the receiver provides for a 3D-engine that takes care of the scaling and inclination of the main TV programme, once the signer video is overlaid. The tilt angle could be fixed or could be controlled by the end-user, for example by means of an IBB app.

In principle, the signer video stream can also be displayed on a second screen, which would then have to be positioned next to the TV set in order to avoid visual refocusing by the consumer when watching the TV programme. Successful tests have already been carried out at IRT with a prototype TV-Set being capable of HbbTV 2.0.1. According to information from a German association for the deaf and hard of hearing, this scenario can be well envisaged in case a second video decoder is not available in the main TV set or set-top box. In this case, the main TV screen remains unaltered in size and inclination. This option might also be interest in public places to allow identification of the sign-language interpreter from relatively far distances (on relatively large second screens).

---

[4] For live TV broadcasting, delay in signer video is natural but may, in principle, also be compensated by delaying the main broadcast signal. On the other hand, sign-language users are used to some kind of delay in live interpretations.

3.1.3 Technical realisation of a virtual signer using video glasses

Alternatively to second screens such as tablet PC or PC monitor, the use of video glasses may be an attractive solution for the reproduction of the TV image with the superimposed signer or the signer video only. As the signer video would be synchronized with the main TV programme, the sound of the main TV programme would be in sync with both TV signals. But it is true to state that more tests and some technical developments need to be performed before such solutions could be considered operational.

At the International Radio Fair 2017 (Berlin, 1-6 September 2017), the German public service broadcasters demonstrated the technical possibility of using video glasses as display mechanism. Figure 5 depicts such a headset including glasses for virtual displays.

For the demonstration, the Microsoft "Hololens" was used. For this demonstration, a locally stored broadcast signal was used while the signer video was stored in the Hololens (as there was the risk that the WLAN at IFA would not be reliable enough for a continuous video stream). The initiation of the signer video was triggered by an HbbTV application. The synchronisation mechanism of HbbTV 2.0.1 [1] was used to time-synchronise the broadcast and the broadband signal (inter-device synchronisation using MPEG TEMI[5] timeline). Figure 6 shows a view through the Hololens depicting the virtual image of the signer next to the main TV screen.

This device configuration demonstrated the feasibility of combining a conventional broadcast TV signal on air with an IP stream of the signer video controlled via HbbTV and delivered via the Internet.

FIGURE 5

The video glasses used in the demonstration at IFA Berlin 2017

Report BT.2448-05

5 Timed External Media Information.

FIGURE 6

Photograph taken through the video glass (Hololens), showing the virtual signer video "next" to the main screen (note the edge of the right glass)

Report BT.2448-06

The showing at IFA 2017 demonstrated that video glasses are, in principle, suited to display a virtual image of the signer video. The focus can be adjusted such that the image appears in the focal plane of the main TV screen. This reduces eye fatigue. More tests are now necessary with people who depend on the sign language for their consumption of audio-visual content.

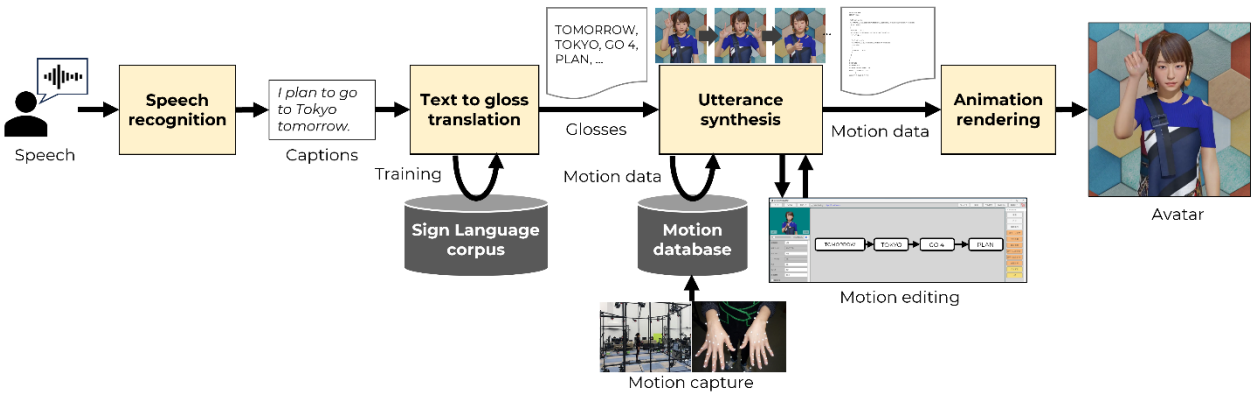
A non-neglectable advantage of video glasses is that the virtual image is only visible to those wearing these glasses und would thus not be recognisable by other viewers that may be disturbed in case the signer is superimposed to the main TV programme. On the other hand, it may not be too comfortable wearing the video glasses for an extended period of time. In future, however, there should be models of more comfort (less weight).

3.2 Generation of avatars

Avatar sign language interpreters may be generated from spoken language through a series of translation processes, e.g. speech → captions → glosses → motion data → avatar as shown in Fig.7.

FIGURE 7

Generation of avatars sign language avatars



(1) Translation from speech to captions

In this process, a speech recognition module converts input speech into captions using speech processing.

(2) Translation from captions to glosses

In this process, a text-to-gloss translation module, which uses a neural machine translation model pre-trained on a sign language corpus for example, converts input captions into a sequence of glosses (textual representation of sign language – signs, expressions, and concepts – through written text).

(3) Translation from glosses to motion data

In this process, an utterance synthesis module converts input gloss sequence into motion data using multiple data read from the sign language motion database.

(4) Translation from motion data to avatar

In this process, an animation rendering module converts sign language motion data into sign language avatar animation by controlling the avatar using the input motion data.

3.3 Generation of avatars at receiver side

Generation of avatars at the receiver side can significantly reduce the transmission data compared to delivery of video streams of signers. It is also possible to customize the size and position of the main video and sign language windows and to personalize the avatar's appearance.

There are three possible strategies to transmit sign language information over-the-air (OTA) or over-the-top (OTT) for controlling an avatar rendered at the receiver side (or companion device):

– Closed caption streaming for automatic translation and adaptation to sign language
  - In this strategy, an application running on a TV receiver or a companion device can request access to the closed caption stream from spoken languages through an API mechanism and automatically translate the caption content into sign language.

– Sign language gloss stream
  - In this strategy, sign language content is represented using a sequence of glosses (textual representation of sign language – signs, expressions, and concepts - through written text). The player decodes the sign language files at the receiver and then synchronizes and renders the sign language content using an avatar.

– Sign language motion stream
  - In this strategy, sign language content is represented by motion blocks coded over time. The motion blocks transmitted are decoded and presented in the receiver through an avatar animation engine, which interprets and uses them to move the avatar and show the signs.

The avatar that performs sign language interpretation is modelled in a humanoid shape. It also assumes that both body and face need to be animated to interpret sign language efficiently. Body and face use different approaches for animation. Therefore, body and face motion shall be treated separately.

The body animation adopts an approach based on the movement of the joints of a skeleton. The skeleton is described as a set of joints. By analogy, each joint corresponds to the connection between the bones of a human body. The skeleton is represented in a tree structure in which, starting from a root joint, new joints are created as children of a parent joint. The root joint has no parent, and all other joints are its descendants.

The avatar used for sign language is presented only in the upper body (head, trunk, upper members), and the root joint is related to the body hip.

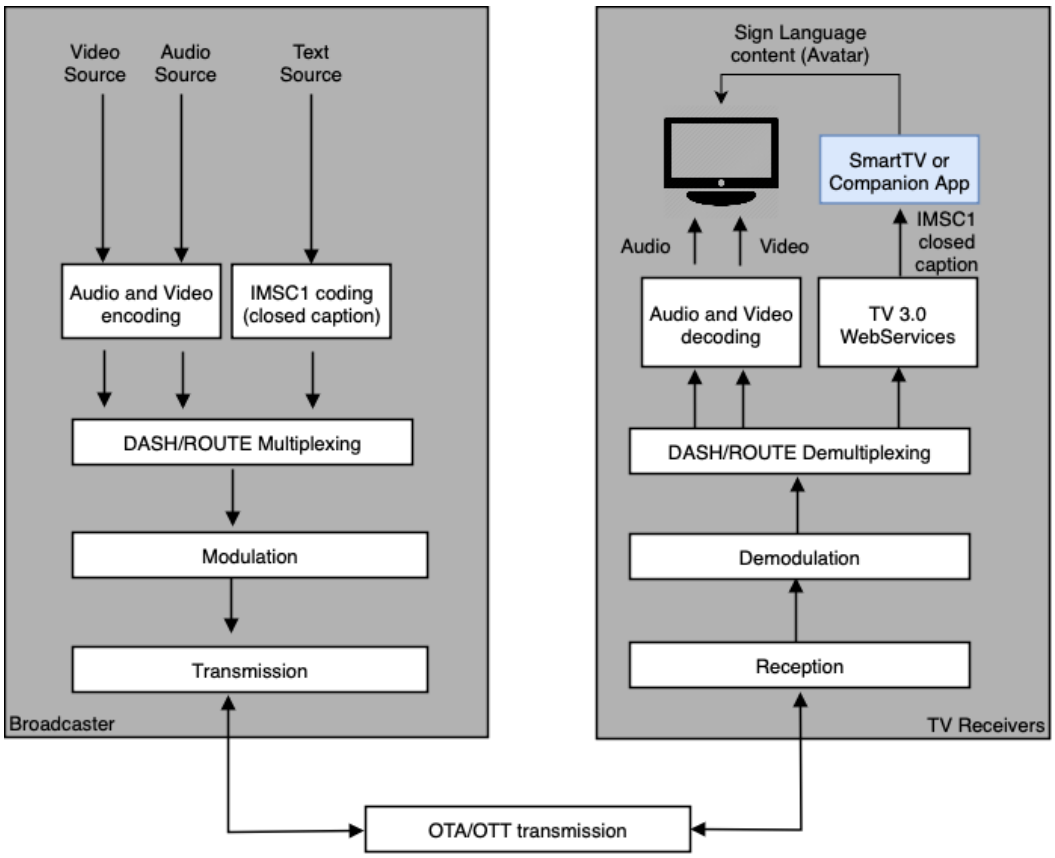
A skin may be drawn around the skeleton so that the visual appearance of the avatar is customized in the receiver. The face meshes are attached to the head-related joints.

3.3.1 Closed caption streaming for automatic translation and adaptation to sign language

This technology enables an application running on a TV receiver or a companion device to access the closed caption stream from spoken languages through an API mechanism, to automatically translate the caption content into sign languages. Figure 8 shows a schematic view of this technology.

FIGURE 8

Schematic view of closed caption streaming for automatic translation and adaptation to sign language technology



The technology works as follows. The broadcaster multiplexes and transmits audio, video and closed caption in written language via OTT or OTA. Then, a registered application requests access to the closed caption stream files in written language through an API mechanism and automatically translates the content into sign language.

In this technology, the sign language machine translation, the sign dictionary and the sign player components are embedded and executed within the application running on a TV receiver or a companion device.

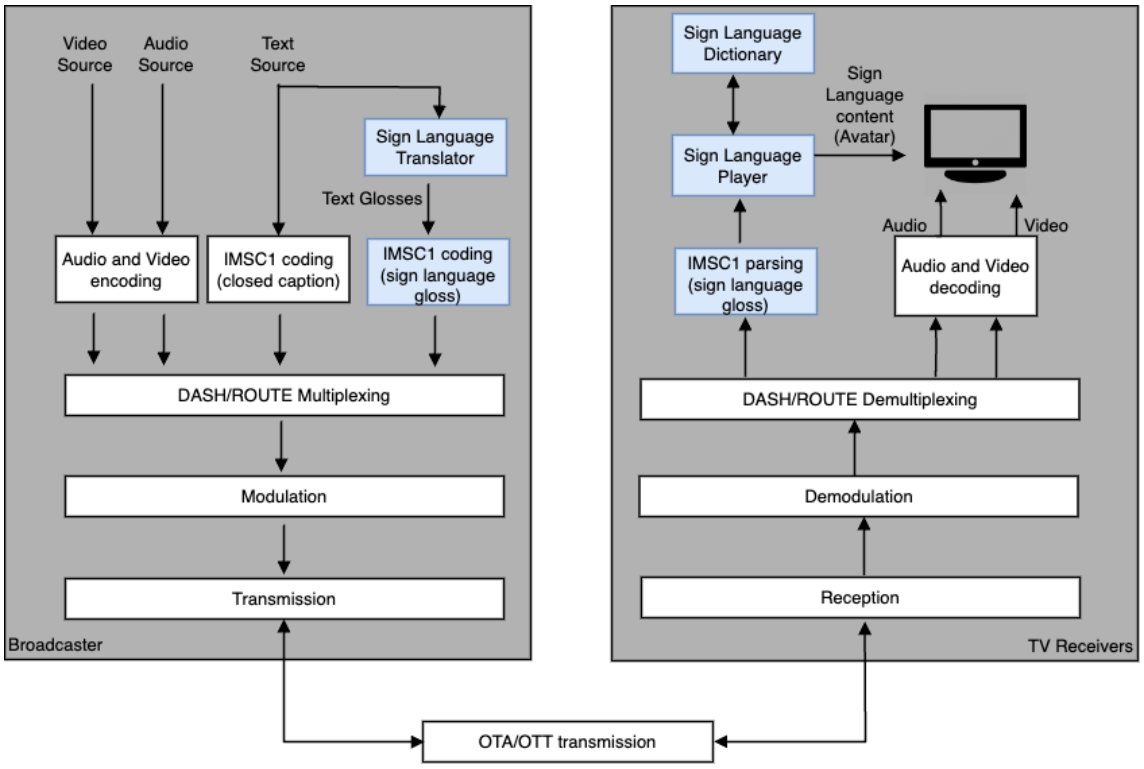
Using the closed caption streaming for automatic translation and adaptation to sign language requires no dedicated sign language production on the broadcaster side, as the closed caption production is reused. However, on the receiver or companion device side, there is all the complexity of translation, adaptation and rendering in real time and with low latency.

3.3.2 Sign language gloss stream

This technology enables the transmission of the sign language content using a sequence of sign language glosses (textual representation of sign language, i.e. signs, expressions and concepts, through written text). The sequence of glosses is embedded in IMSC1 files and transmitted in the DASH stream. At the receiver, the player decodes the IMSC1 sign language gloss, synchronizes and renders the sign language content using an avatar. Figure 9 shows a schematic view of the sign language gloss stream technology.

FIGURE 9

Schematic view of the sign language gloss stream technology



The technology works as follows. At the broadcaster facilities, a sign language machine translation component receives the text or closed caption stream of the TV content in the spoken language and translates this content into a sequence of glosses in sign language. Afterwards, it encapsulates the sequence of glosses in IMSC1 files. Then, the IMSC1 files containing the sign language glosses are multiplexed into the DASH stream and transmitted to the receiver. The sign language gloss can be transmitted via OTT or OTA.

When the user enables the sign language option at the receiver, the sign language player receives the IMSC1 files of the gloss stream, decodes these files, and renders (or synthesizes) a sign language window with the support of a sign language dictionary. The sign language dictionary shall contain a set of animations for the sign language glosses.

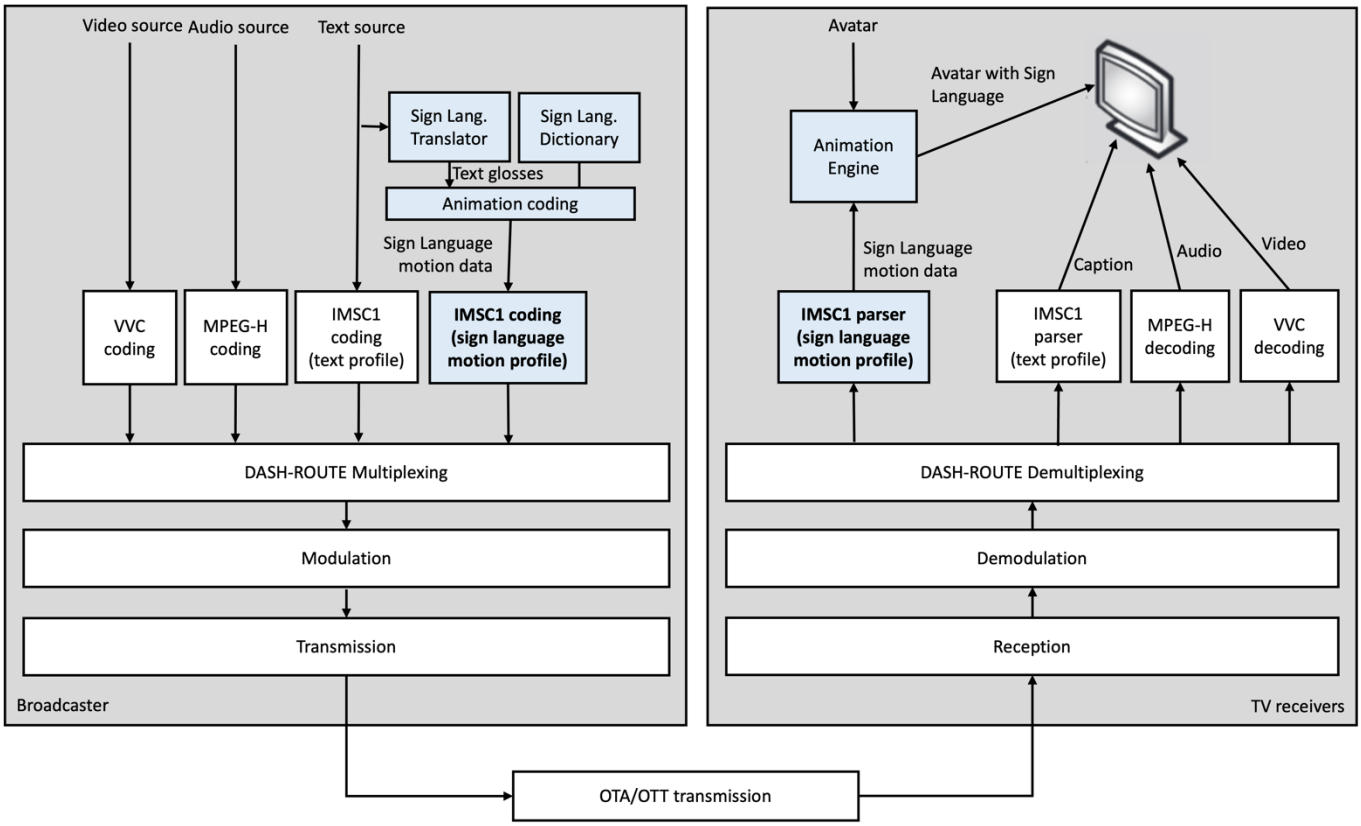
Using the signal language gloss stream requires the translation from spoken language to gloss on the broadcaster side and requires an animation dictionary to render the gloss as sign language on the receiver or companion device side.

3.3.3 Sign language motion stream

Figure 10 exemplifies the process of transmitting a sign language motion stream.

FIGURE 10

Transmission of sign language motion stream



From the broadcaster side, a sign language translator defines a sequence of glosses to interpret the audio/video context in sign language. A sign language dictionary can be used to create a file containing information about the movement corresponding to that sequence of glosses. This file is coded in IMSC1 format and multiplexed with audio, video and closed captions to be transmitted via OTA or OTT.

From the TV receiver side, content from the broadcaster is demultiplexed, and a motion file is extracted from IMSC1 content. The face and body movement data stored in the motion file feeds an avatar animation engine, which animates a predefined avatar according to that data. The result is the avatar moving its body and face to interpret the audio/video context in sign language.

Using sign language motion stream requires that all the translation process from the spoken language all the way to the motion blocks be done on the broadcaster side, but the processing requirement on the receiver or companion device side is lighter.

3.4 Sign-language Computer Graphic system using mobile devices in Hybridcast service

At the occurrence of natural or man-made disasters, people seek to be informed through television and radio broadcasting to understand what is happening or about to happen. It is desirable for peoples with hearing disabilities to receive information by means of sign language, but signers are not always available for broadcasting.

A prototype of a sign language Computer Graphic (CG) presentation system has been developed experimentally [6]. The system automatically generates sign language CG animation corresponding to emergency calls made by announcers upon receipt of, for example, earthquake early warnings. The system uses a Hybridcast function that enables collaboration of companion devices to present CG animation of a signer on a companion device as a secondary screen such as smartphones and tablets linked to the TV set (see Fig. 11).

During a Hybridcast service being received on a TV set, when an Event Message for the emergency alert is received, which is a trigger signal multiplexed in a broadcasting signal, the TV set automatically turns off the Hybridcast content currently being overlaid on the TV programme and switches to the emergency content. Using the function of Event Message, the sign language CG system can automatically present an emergency information by the sign language CG animation on the secondary screen linked to the TV set. It also enables sign language information to be presented only to the secondary screen devices of those needing the information during the occurrence of a disaster without interrupting the broadcast programme.

An application on the secondary screen device, which generates the CG Animation and subtitles, requests the delivery server for TVML (TV program Making Language) [7] and subtitle data. The server then sends the requested data back to the application. Using the location information such as postal code, prefecture code and the area code, stored in NVRAM (Non-Volatile RAM) of the TV set enables the provision of dedicated information such as evacuation information and shelter location announced by local governments, corresponding to the user's residential area as shown "Area A" or "Area B" in Fig. 11.

The system adopted the client-side rendering using the WebGL based TVML player display the CG animation onto a secondary screen device (see Fig. 12). TVML player is a CG animation creating system, and WebGL is a 3D graphics API that is mounted in standard web browsers. By using the WebGL based TVML player, it is possible to draw the CG animation on the web browser without installing specialized software in end-user devices.

FIGURE 11

Structure of sign language CG presentation system using a companion device

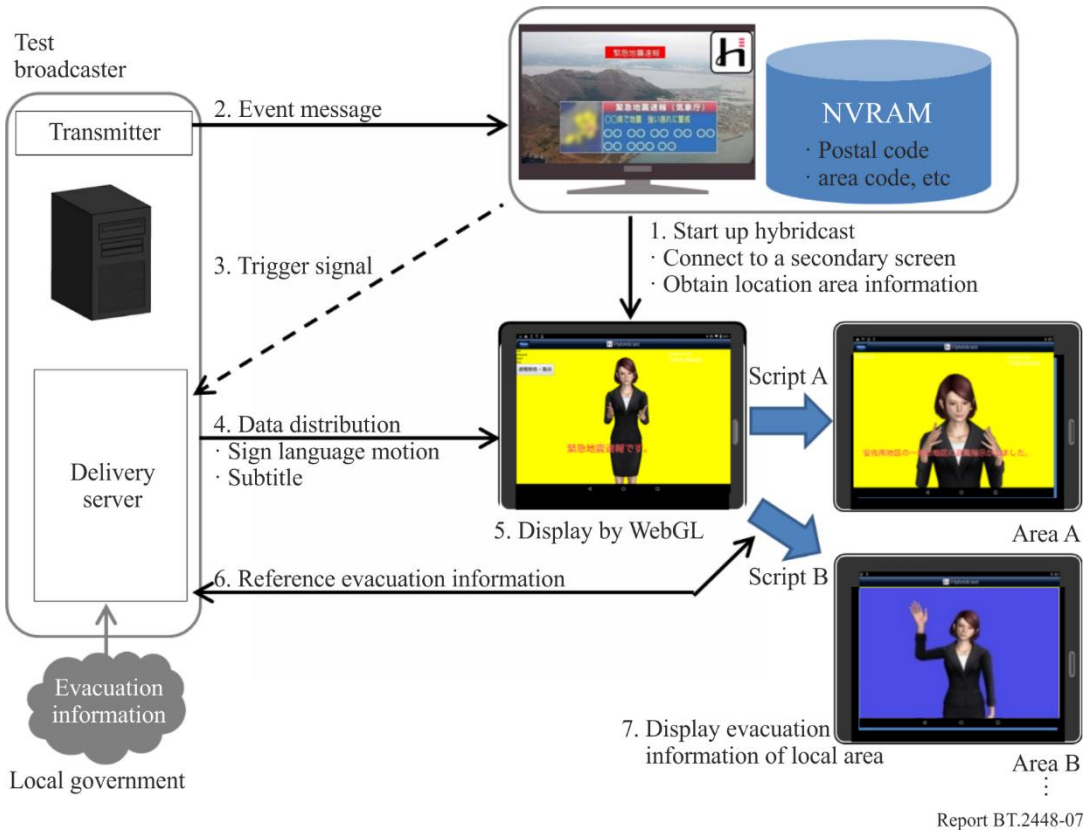
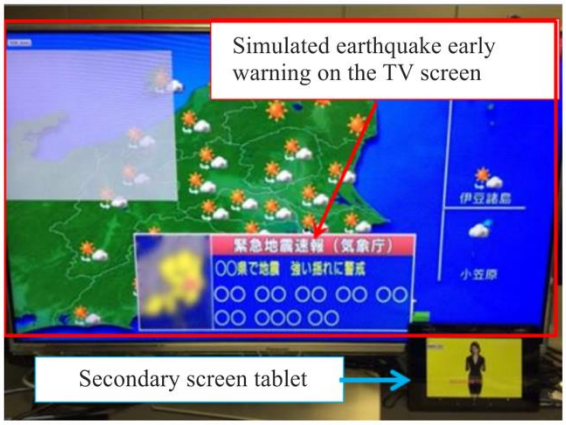


FIGURE 12

Prototype system

a) Hybridcast compatible TV and Tablet



b) Examples of displaying Japanese Sign Language CG animation onto a companion device

Report BT.2448-08

3.5 Single screen closed signing using IBB

This section describes the case that signer video is displayed on a TV screen together with a TV programme on an IBB receiver without a secondary screen device. The signer video can be a CG animation as described in § 3.4. High performance is required for the execution and rendering of graphics in order to generate CG animations, which is not practical for IBB commercial receivers. An alternative method is to generate signing video images at the server side and distribute them to the receivers as normal video streaming. In this case, an IBB receiver needs two video decoders for displaying the CG signer video on a TV programme. Figure 13 shows an implementation of a single-screen closed signing for a practical IBB receiver.

At first, the IBB receiver requests an application for closed signing presentation and acquires the application. Once the application starts on the receiver, it requests the video stream of the signing CG to a server. Video images for closed signing can be generated on the server by using CG, either in real time or in pre-recorded form. Real time rendering is useful for providing a closed-signing service even with live broadcast programmes. The method described in § 3.4 can be used to generate closed signing video images, which involve TVML for CG animation. The rendering server gets the TVML data from the broadcaster. The rendering server parses the TVML data and renders CG animation on a TVML player. The rendering server captures the rendered CG animation simultaneously and transcodes it to MPEG-DASH video segment files. The IBB receiver gets video streaming by using MPEG-DASH, and decodes it using one of the two decoders while the other video decoder decodes the broadcasting video. The application controls composition of these two video images for a single screen in accordance with the users' manipulation.

Figure 14 shows an example of displaying sign language CG animation with a prototype Hybridcast receiver which has two video decoders to display both signing streaming and broadcasting video.

FIGURE 13

Implementation of single screen closed signing for a practical IBB receiver

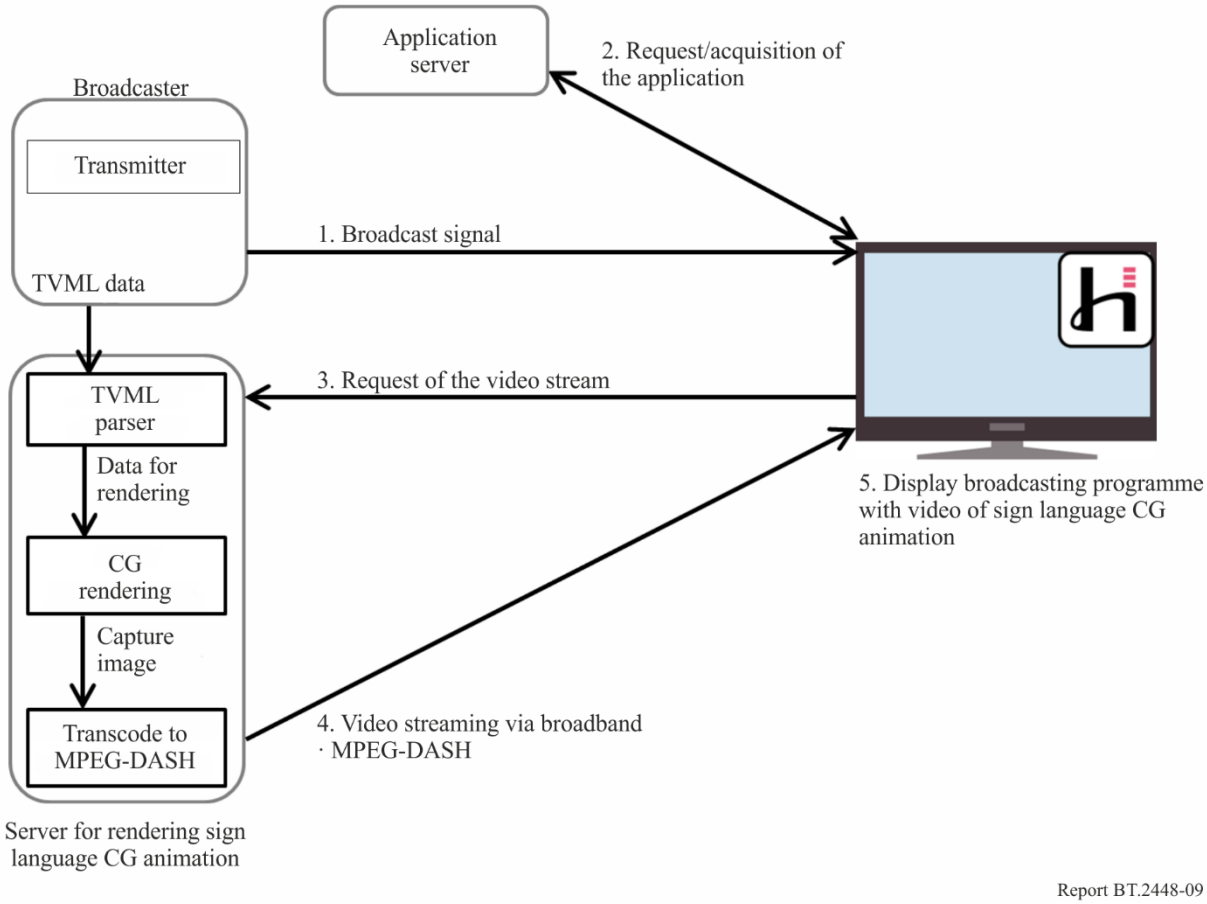
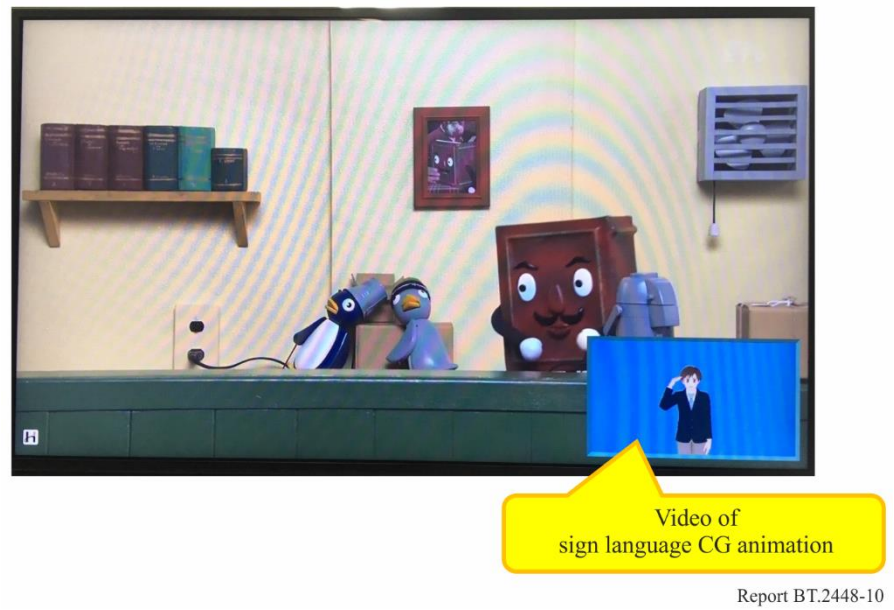


FIGURE 14

An Example of displaying sign language CG animation on Hybridcast compatible TV



3.6 Conclusions on methods of providing closed signing

Solutions for closed signing in legacy TV receivers relate to selection concepts, i.e. the consumers can select between the main TV programme without inserted sign-language interpreter, and the main TV programme with superimposed signer video. The composition is accomplished in the studio. The main TV programme with the inserted signer is either transmitted in the broadcast multiplex or offered via the broadband Internet. Several versions of the TV signal with the inserted signer can be prepared for selection by the end-users.

Signers can be either real persons or CG avatars. Generation of avatars require a series of translation processes from spoken language to captions, glosses, motion data, and avatars. Using avatars do not require a human sign language interpreter for the programme production, facilitating the automation of such production.

For more advanced IBB TV sets or set-top boxes that are equipped with a second video decoder, the transmission of the signer only (related to a given TV programme) is possible. This saves transmission capacity and allows for almost unlimited setting of user preferences with respect to size and position of the signer video as well as with respect to the display of the main TV programme (size, inclination). Alternatively, video glasses may be used to reproduce the signer video as a virtual image next to the main TV set.

Transmitting sign language information for controlling an avatar rendered at the receiver side (or companion device) saves even more transmission capacity. The three different strategies available for this kind of transmission present different levels of complexity on the broadcaster side and on the receiver or companion device side.

References

[1] Recommendation ITU-R BT.2075 – Integrated broadcast-broadband system

[2] Report ITU-R BT.2267 – Integrated broadcast-broadband systems

[3] D6.1 – Pilot-D Progress report, HBB4ALL deliverable, December 2014. (http://www.hbb4all.eu/wp-content/uploads/2015/03/D6.1-Pilot-D-Progress-Report.pdf).

[4] D6.2 – Pilot-D Solution Integration and Trials, HBB4ALL deliverable, October 2015. (http://www.hbb4all.eu/wp-content/uploads/2015/03/D6.2-Pilot-D-Solution-Integration-and-Trials-2015.pdf).

[5] D6.4 – Pilot-D Evaluation and Recommendations, HBB4ALL deliverable, Dec. 2016. (http://www.hbb4all.eu/wp-content/uploads/2017/02/D6.4-UPM_Pilot-D-Evaluation-and-Recommendations_v1.00.pdf).

[6] T. Uchida, S. Umeda, M. Azuma, T. Miyazaki, N. Kato, N. Hiruma, S. Inoue, Y. Nagashima, “Provision of Emergency Information in Sign Language CG Animation over Integrated Broadcast-Broadband System”, IEEE BMSB 2016, 11C-3, (2016).

[7] H. Kaneko, N. Hamaguchi, M. Doke, S. Inoue, “Sign Language Animation Using TVML”, ACM VRCAI 2010, p.289-292, (2010).

Annex

Representation of the signer video on the main TV screen (example Japan)

1 Introduction

The size and position of a signer image are the important parameters in not only the open signing but also closed signing. In this Annex, some information is provided on how to superimpose the signer video with the main TV programme.

This Annex provides information on how signers are presented in current TV programmes with signing (open signing) in Japan.

Signing is the first language of many people born without the ability to hear and is a unique language. Captioning alone is not enough since signing is easier for the hearing impaired to understand.

Japanese broadcasters provide numerous programmes with closed captioning but few with signing. The information reported in this Annex is based on five programmes broadcast by NHK, Nippon TV, TBS, TV Asahi and Fuji Television.

2 Size and position of a signer

The broadcasters are aware of the UK OFCOM guidelines "Ofcom's Code on Television Access Services", which state that the image of the signer superimposed upon the original programme should generally appear on the right hand of the screen and occupy a space no smaller than one sixth of the picture.

In Japan, there are no common guidelines concerning how a signer should be presented in TV programmes. Signers are positioned and sized differently in different programmes as shown in Fig. 15. There are two types of programmes with signing: one is produced specially for the hearing impaired where a signer is the main presenter, and the other is produced by synthesizing a signer image and the normal images.

Programme (a) is a news programme dedicated to the hearing impaired in which the signer is placed in the centre and subtitles are shown on the right-hand side of the screen.

Programmes (b) through (e) show the singer in the lower-right of the screen. Two types of synthesizing methods are used: chroma key and small circular window ("wipe"). When the wipe method is used, the window takes up half the screen in the vertical direction.

In programme (b), the positions of the main presenter and subtitles are shifted left to avoid interference with the signer image.

Programme (c) uses the chroma key method, and the signer is shown a bit larger depending on the feedback from the audience.

In programme (d), the size and position of the signer are fixed because moving the position of the signer makes it difficult for the audience to see the signer, although the signer and the main presenter or subtitles sometimes overlap.

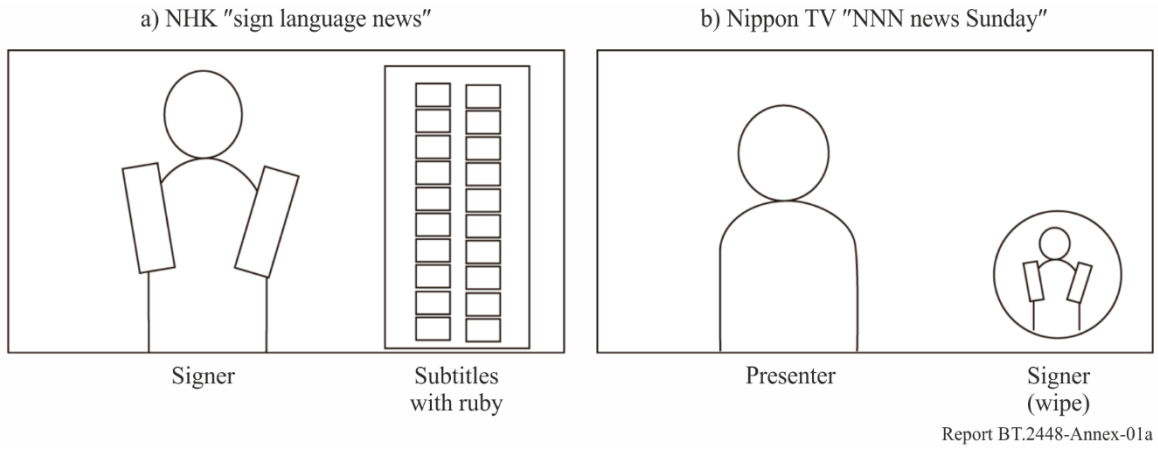
Programme (e) uses a "soft wipe" that blurs the edge of the small window, but the size and position of the signer are more or less the same as in programmes (b) and (d).

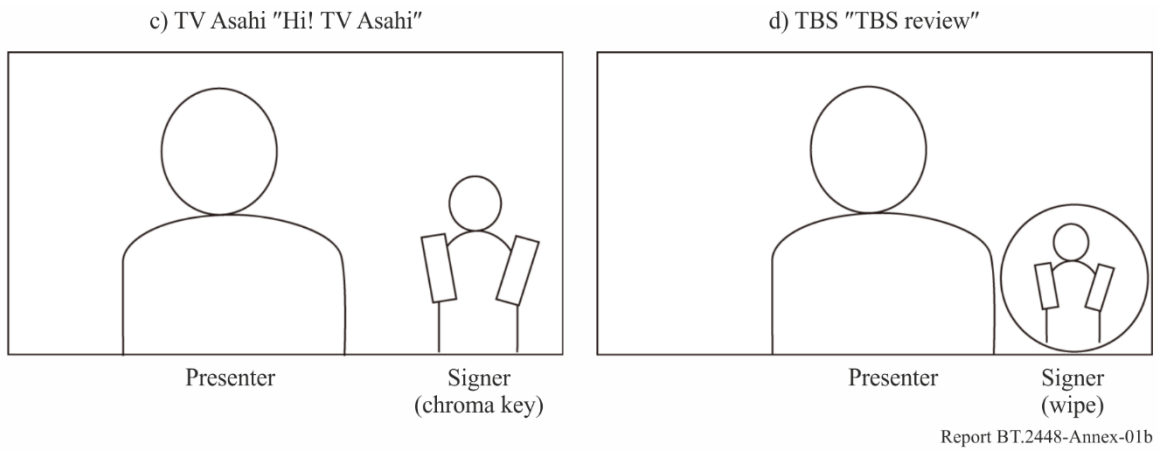
According to the feedback from the audience, hearing-impaired viewers do not prefer the small-window method because sign-language expressions, in particular facial expressions, should be easy to see. In closed signing, studies are required on how best a signer image should be presented

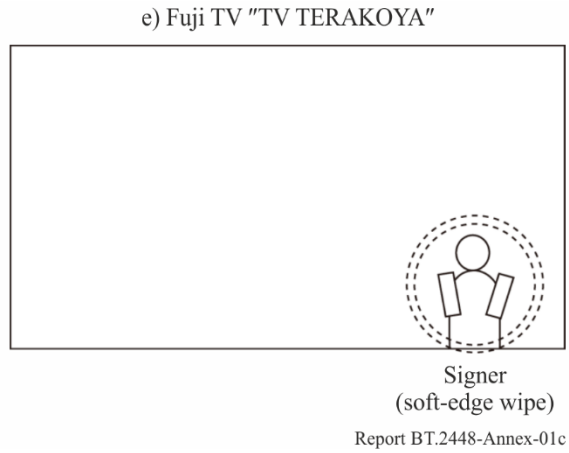
with the main programme in a receiver considering the opinions of the hearing impaired as well as those of the programme producer.

FIGURE 15

Size and position of signer in frame







3 Frame rate to reproduce signer's motion

The signer image has the same 30 frames per second or 60 fields per second as the normal video, and there have been no complaints that the current frame rate is insufficient for viewing sign language. A study would be required to determine the minimum required frame rate for an easy-to-understand sign-language video in closed signing.

4 Speed and synchronization of signing with main video

The speed of sign-language is usually adjusted to the aural commentary, but when complicated sign language expressions are required, a delay between the signing and the commentary may occur. To prevent any delay, the lengths of the comments for the sign language must be adjusted and care given to the switching timing of the video sources. In the closed signing, a synchronizing mechanism would be needed.

5 Example of signing information service through the Internet

Prior to the deployment of TV receivers for closed signing, NHK released a signing information service about weather forecast through the Internet (http://www.nhk.or.jp/strl/sl-weather/) (see Fig. 16). On the website, video showing the latest weather information of a selected area including weather marks, temperature, and subtitles together with a CG signer animation is provided in accordance with the request from users using a PC or a smartphone (see Fig. 17).

The CG signer animation is automatically generated on the basis of meteorological data distributed from the Japan Meteorological Agency at least three times a day. The various pieces of information are laid out for ease of viewing.

As the next step, a service to provide push notifications of emergency alerts is under consideration.

FIGURE 16

Top page of signing information site about weather

Report BT.2448-Annex-02

FIGURE 17

Layout of CG signer and weather information

Report BT.2448-Annex-03

6 Conclusion

The current status of TV programmes with signing in Japan is reported. The size and position of a signer image are the important parameters in not only the open signing but also closed signing. It is hoped that this information will be of some help in the study of the requirements of TV receivers for closed signing.