ATVOD Working Group on Access Services (WGAS)

2015 Report   prepared by Nick Tanton

Introduction

Our first report delivered in 2014 discussed the provision of subtitles and audio description with on-demand content but focussed specifically on the issues and challenges associated with delivering subtitles to the OD consumer.

That report

· observed that, whilst the technical obstacles are straightforward and surmountable, delivery of accessible OD content often involves commercial challenges,

· proposed measures for facilitating the transfer of access-services assets and

· recommended that EBU-TT (or variants thereof) be used as the format for subtitling authoring, exchange, archiving and, where possible, delivery.

This 2015 report first concentrates on Audio Description [1] for Blind and Partially Sighted people (B&PS). It then provides an update on progress in subtitling.

As before, the content of this report represents a distillation of understanding from within the ATVOD Working Group on Access Services (WGAS).

Audio Description

Audio Description (AD) is a service that offers an alternative version of the programme sound into which short descriptions of salient scene detail or action have been inserted into gaps in the original narrative for the particular benefit of B&PS viewers who might otherwise miss purely visual cues.

Any perceived barriers to distributing AD with on-demand content often invoke the difference between the two methods that are used in the UK on linear platforms to delivering AD to the home (i.e. on the final leg of the journey). These are known as "broadcast-mix" and "receiver-mix" AD [2]. However most if not all of the steps involved in authoring, playout and manipulating AD use the same data structures and signal forms up to the point of coding for delivery. Only at this point is the AD put into a form suitable for the particular delivery platform and for the capabilities of the user terminal.

---

[1]   or "Video Description" as it is called in the US

[2]   With "broadcast-mix" AD (referred to as "B-mix AD"), programme sound and description are combined before delivery to the end-user-equipment. By contrast, "receiver-mix" AD (a.k.a. "RX-mix AD") involves mixing decoded programme sound and decoded description in the end-user-equipment.

Authoring, Content Exchange and Versioning AD

The process of authoring audio description in practice only starts when the programme material has been edited into its "transmission-ready" form. Until then it won't be fully clear what opportunities there will be for description (i.e. sufficiently long gaps in the narrative) or exactly what needs to be described. Once edited and ready for description, a trained describer will then review the finished content, produce a description script with timings and fade levels and then either record the descriptions to that prescription or pass the script and timings to someone else to voice the AD accordingly.

Importing content which has already been described (e.g. for linear television delivery or for cinema presentation) will usually require a number of subsequent processes. If the imported content is itself "hard-parted" (partitioned episodically e.g. for advert breaks etc.), video editing will be required to construct a continuous narrative for any different version. If the description is available as a separate file or "track", there may be words used in it which have different meanings (elevator/lift, sidewalk/pavement etc.) [3]. The content may have a different frame-rate from that to be used for the OD service, in which case time-codes need to be converted [4]. Short of re-voicing the description completely, re-versioning may involve dropping entire original description passages around any new edit points [5]. If the description is already a mix of narrative, effects and description, then editing the AD as supplied is infeasible and re-voicing the new edition is the only practicable option. For imported content therefore any AD processing will be potentially minimised if content exchange involves a separate clean description.

If the content is indeed exchanged with separate description (i.e. not already mixed with narrative, music and effects) the description data should contain audio for each description passage, timing (in/out or in/duration) for that passage and ideally a fade-value to be applied to the contemporary music and effects (M&E) during that passage so that the description can be heard above M&E [6]. Since the late 1990's UK broadcasters have used an additional linear audio track associated with the video and with the programme sound to author, cache and convey this description through playout to coding prior to delivery to DSAT, DTT and DCable. The relevant signal format is widely used and is described in http://www.bbc.co.uk/rd/publications/whitepaper198. There are, in principle, mechanisms that could be used to embed and convey this data within EBU-TT [7] but any initiatives to further develop this have yet to take place.

---

3   If a difference in meaning is the *only* reason for editing then it is generally recommended *not* to intervene

4   This is a process which is, under some circumstances, amenable to automatic tools.

5   with a possible loss of service quality for the AD user

6   Alternatively a default fade value might be applied for all description passages.

7   A move to the use of EBU-TT for subtitling was recommended in the WGAS 2014 Report.

For short-form content (trails etc.) the narrative is often almost continuous and so typically provides little opportunity for description.

One issue that usually needs clarification when described content is exchanged is to establish on which audio tracks are to be found the description and attendant control data (or pre-mixed narrative, effects and description). That may depend on the content provider and whether the content is archived standard definition or HD material. EBU document R 123 [8] describes the track numbering for Broadcast Content Interchange including Audio Description.

The Digital Production Partnership (DPP) technical standards [9] for delivery to UK broadcasters includes a paragraph on additional audio-only files (such as AD) which prescribes the use of BWF (aka "B-WAV" files) conforming with EBU-Tech 3285, with file duration and timecode exactly matching the principal MXF file.

Delivery of AD

The delivery of "broadcast-mix" AD to the home terminal requires only that the pre-mixed AD stream can be explicitly selected in the home terminal (e.g. as a labelled additional audio track). "Receiver-mix", on the other hand, requires that the home terminal mix the "programme sound" (narrative and effects) with the description under the control of the accompanying fade values.

"Broadcast-mix" AD precludes the user being able to adjust the relative volume of the description compared with narrative and effects. This can be a problem as not all describers will be equally audible to all users, especially as some AD users may also have aged-related hearing loss. "Broadcast-mix" AD uses somewhat more bandwidth (to convey music and effects as well as voice). Furthermore, as a B&PS user will find it particularly hard manually to change audio from main to described stream and back, the end-user-equipment needs to manage how to present material that doesn't for some reason have description (e.g. offer a fall-back mode to standard audio).

"Receiver-mix" AD does allow the user to adjust the relative mix of description and narrative/effects to suit their taste and capabilities which results in a better user experience. The RNIB view is therefore that, on balance, "receiver-mix" AD is the delivery method to be preferred when feasible. For receiver-mix AD, the description component need only convey voice and need not be present as a stream-component on programme content that is not described.

---

[8] https://tech.ebu.ch/docs/r/r123.pdf

[9] A generic version of those standards is available at http://www.digitalproductionpartnership.co.uk/what-we-do/technical-standards/delivery-standards.

As noted above AD is delivered to UK linear digital platforms in either form and the major channel providers have established automatic mechanisms to manage AD to acceptable levels of reliability as perceived by the user.

ATVOD WGAS is of the view that if a platform is physically capable of supporting RX-mix AD it should do so, but that it would be better to deliver B-mix AD than nothing at all.

ATVOD's 2014 survey of access service provision [10] identifies services provided with AD at that time. It is believed that other VOD platforms have since enabled AD with the potential to add AD categories to content selection. However we note that delivery of AD in a recognisable or easy-to-use form is dependent on the end-user equipment and its software over which providers may not have much control.

Work is continuing elsewhere both on "object-based audio coding" and on "dialogue enhancement". In the longer-term these may have an impact on the opportunities for AD delivery and a watching-brief should be kept on these developments.

Companion devices

The development of "companion devices" has resulted in proposals that specialist services (such as AD) might be delivered specifically to such accessories. DVB has a public specification ETSI TS 103 286 [11] which HbbTV 2.0 references and this would allow applications to offer synchronised ancillary service components on companion devices.

The practicalities of suitable companion devices in the context of B&PS users have yet to be fully explored. The RNIB has been trialling an app which provides AD synchronised with programme sound. In their words: "The trial has shown that audio fingerprinting can be a viable technology for synchronising audio description to a program. The majority of trial participants reported that it made a great difference and offered a welcome service."

A watching brief should be kept on these developments. There is a potential risk to service providers and platform operators if a proliferation of different companion devices were to be used to convey access services such as AD or STs and used in conjunction with the already diverse set of VOD platforms and equipments. This would compound the existing challenges of delivering access services to multiple platforms and to multiple devices.

---

[10] http://www.atvod.co.uk/uploads/files/Provision_of_Access_Services_2014_Report_FINAL.pdf

[11] http://www.etsi.org//standards-search?search=TS103286&page=1&title=1&keywords=1&ed=1&sortby=1

Signalling the availability and presence of AD

Not all OD content will be described (some may even be unsuitable for description) and, because AD is mainly used by the B&PS community, it is important for service providers to include or to reliably support mechanisms that allow a B&PS user to know that AD is actually available for that content. DVB provides simple tools for timely signalisation within DVB-compliant streams (including distinguishing between broadcast-mix and receiver-mix AD stream components) and these can be used or as necessary adapted for other platforms.

Equally any "metadata" or signalisation associated with the content to support such mechanisms must be consistent, accurate and up-to-date (e.g. "this content has/doesn't have description") as provided by the OD service provider. This also applies to managing different versions of the same content (e.g. for different platforms).

Usability issues at the home terminal

As noted in our 2014 Report, the species (and sometimes generational variants) of end-user equipment are diverse and some may not have been designed a priori with accessibility in mind. There is also still a mind-set which presumes that B&PS persons are not likely to watch or wish to watch AV content, however delivered; their particular needs are therefore not always addressed in the design of the user interface.

Distinguishable audible cues (ranging from tone beeps to voice-synthesis) to complement on-screen messages have been successfully included in various implementations of DTV home terminals in recent years. The UK Switchover Help Scheme, RNIB, DTG, Digital Europe and various TV manufacturers have taken significant roles in establishing, recording and implementing the principles of such B&PS person oriented user-interfaces.

Careful consideration needs to be given both by the platform operator and by the user-interface designer as to how a B&PS user "discovers" on-demand content that is available with description, how they might choose to select (and de-select) described content and how to manage a mixed audience where one person wishes to hear the described version and another in the same room doesn't. Storing and recalling user-preferences in a simple manner suitable for B&PS persons is also important.

AD on linear channels

Ofcom figures showing the growth of access service provision on linear services over the ten years 2004 – 2014 can be found at

http://stakeholders.ofcom.org.uk/binaries/research/tv-research/access-service-reports/Access_service_provision_2004_to_2014-_Final.pdf

Update on Subtitling

1   NBC Universal has undertaken a major tech refresh, and is now using EBU-TT. Working with Screen Subtitling Systems and others, they have developed non-proprietary solutions which will therefore be available for others to use. NBC Universal selected the EBU-TT as their Master File Format for subtitling, as it is an open standard, supports the widest range of foreign language characters & glyphs, is human readable and can be easily manipulated within their automated media workflow.

2   The Current YouView spec has a profile version of TTML which has a lot of overlap with EBU-TT-D.

3   DVB-DASH (aka MPEG-DASH Profile for Transport of ISO BMFF Based DVB Services over IP Based Network) references EBU-TT and ETSI has just published it as technical standard TS 103 285 v1.1.1.

4   A set of test suites for EBU-TT-D will become available to test implementations and support for presenting EBU-TT-D. This may also cover EBU-TT over DASH.

5   The EBU has published a v0.8 draft of the EBU-TT Part 3 live contribution spec (for delivering subtitles from live author or pre-prepared to an encoder) for comments [12].

6   W3C's profile of Timed Text (Internet Media Subtitles and Captions 1.0 (IMSC1)) is currently at the "candidate recommendation stage". There is a finalised test suite for those features which differ from TTML. W3C has also issued a "call for implementations" [13] - anyone implementing such features is invited to inform W3C and submit any tests that they have passed. BBC has contributed a couple of passed tests.

7   W3C's Timed Text Working Group has also published a first public working draft of TTML2, the first public working draft of WebVTT and is working on mapping between TTML and WebVTT.

---

[12]   https://tech.ebu.ch/news/2015/06/19/ebu-tt-subtitling-goes-live

[13]   https://lists.w3.org/Archives/Public/public-texttracks/2014Dec/0000.html

8   The German Institut für Rundfunktechnik (IRT) is working on open-source subtitle format conversion and are seeking to identify essential subtitling features which should be preserved across such conversions. ATVOD WGAS could usefully assist by identifying those features relevant to its members' activities.

Current Understanding of Platform Capabilities re. Subs & AD

Over the past year WGAS has discussed our current understanding of the availability of subtitles and AD on different on-line platforms and for our own purposes have accumulated this understanding in a summary table for our own reference. In some instances this information has been verified, in others it awaits validation. This table is attached to WGAS minutes and will be kept regularly updated.

Conclusions and recommendations

Progress is being made in the provision of Subtitling and Audio Description with on-demand content.

The international momentum behind timed-text is clearly growing. As various organisations re-engineer their internal infrastructure and as equipment manufacturers begin to offer equipment to support EBU-TT (and its conversion to and from legacy subtitle data formats) the technical barriers to subtitle provision described in our 2014 Report will continue to diminish.

1   ATVOD WGAS should, with OFCOM, identify those particular subtitling functionalities and features which it is important to preserve when converting from one subtitle format to another.

The barriers to wider provision of AD currently include some confusion about the format of described content delivery and an absence of understanding or suitable technical functionality in the design of end-user equipment. As noted above providers may not always have much influence on end-user equipment/software.

2   We recommend that VOD service providers and content aggregators seek actively to acquire content with audio description when description is already available and when any target delivery platform or end-user equipment can support AD in some form.

3   Content providers should be strongly encouraged to offer description as a separate track or file regardless of whether the delivery platform or end-user equipment supports B-mix or RX-mix AD. This also makes any post-acquisition processing possible.

4   ATVOD WGAS recommends that, if a platform is physically capable of supporting RX-mix AD, it should do so. It nevertheless recognises that it is better to deliver B-mix AD than no AD at all.

5   ATVOD should take its part in encouraging service providers and platform operators to consider all the features of service discovery, ease-of-use etc. that will make it straightforward for a B&PS person to access audio-described on-demand content.

6   ATVOD should track developments in "companion devices" and understand the wider implications for delivering accessible on-demand content which such devices.

List of contributors

ATVOD, BBC, BT, Channel 4, Channel 5, Deluxe, Discovery, Digital TV Group (DTG), ITV, NBC Universal, Ofcom, Red Bee Media/Ericsson, RNIB, Sky, STV, Turner

Nigel Megitt (BBC) as chair of EBU XML Subtitles group and co-chair of W3C Timed Text Working Group

Nick Tanton (consultant) as chairman of ATVOD WGAS

Dr Pablo Romero-Fresco (University of Roehampton)