Standards for Multi-Stream and Multi-Device Media Synchronization

Given the commercial interest in media synchronization and disadvantages of proprietary technologies, consumer-equipment manufacturers, broadcasters, and telecom and cable operators have started developing a new wave of international standards for media synchronization. This article provides an overview of recently published standards from the most relevant bodies: IETF, ETSI, MPEG, DVB, HbbTV and W3C.

M. Oskar van Deventer, Hans Stokking, Matt Hammond, Jean Le Feuvre, and Pablo Cesar

Abstract

Media synchronization is getting renewed attention with ecosystems of connected devices enabling novel media consumption paradigms. Social TV, hybrid TV, and companion screens are examples that are enabling people to consume multiple media streams at multiple devices together. These novel use cases require media synchronization, as unfortunately there are substantial delay differences between the various delivery routes for television and streaming media. Broadcasters have started using proprietary solutions for over-the-top media synchronization, such as media fingerprinting or media watermarking technologies. Given the commercial interest in media synchronization and the disadvantages of proprietary technologies, consumer-equipment manufacturers, broadcasters, and telecom and cable operators have started developing a new wave of international standards for media synchronization. This article provides an overview of recently published standards from the most relevant bodies: IETF, ETSI, MPEG, DVB, HbbTV, and W3C.

Communications Standards

Introduction

Media synchronization is relevant whenever two or more associated media streams are played out together. The classic example is synchronization of audio and video for a television broadcast in order to achieve lip synchronization. More recent examples are social television (TV), hybrid TV and companion screen. Social TV, a.k.a. "watching apart together", has multiple users watching the same TV broadcast while communicating with each other by voice, chat, or other social media. Hybrid TV converges multiple media streams from different channels (broadcast, Internet) into one single TV stream (e.g. broadcast video with subtitles or alternative audio received via the Internet). The companion screen provides user interaction or media on tablet devices associated with a television broadcast (e.g. a play-along quiz, alternative audio, or alternative camera views).

Requirements on synchronicity differ per use case. Social TV is the least demanding case. If there is no audio crosstalk, users won't notice delay differences of less than a second, and often they do not even notice a four-second difference [1]. Hybrid TV is the most strict case, since lip-sync requires audio and video to be synchronized within 40 milliseconds [2]. The companion-screen case would be between those two extremes, adding the challenge of achieving synchronization between two separate devices where communication latency must be compensated for.

Even the least demanding requirement cannot be met by today's media delivery technologies. There can be up to six seconds difference in delivery of a single broadcast channel in a single country via different providers [3]. Transcoding buffers are a major contribution to those delay differences. Transmission delays are also significant. For example a single satellite hop introduces over a quarter of a second delay due to the non-infinite speed of light. Internet delivery using content delivery networks (CDN) is by far the slowest delivery technology. It can easily take 30 seconds to perform all required delivery steps, from transcoding and segmenting to segment buffering at the media player client. A recent test showed a 72 seconds delay between a UK broadcaster's origination of a television channel and its delivery via Internet outside the UK [3].

Broadcasters have started using over-the-top media-synchronization technologies based on audio fingerprinting or audio watermarking for offering synchronized companion-screen content, as these technologies are relatively easy to deploy even in the absence of standards. However, they fail when the audio level is low or if there is background sound in the viewing environment, and considerable confusion may occur when a clip from a program is being reused in another program. Any system must make a compromise between factors such as recognition speed, robustness, and perceptibility of any changes to the audio or ability to discriminate across a volume of audio material [4]. Both fingerprinting and watermarking poorly handle user interactions, such as pause, seek, rewind, and fast forward. Another concern is cost. There is the cost for changing the workflow to get the watermark into the audio of a broadcast before the encoder. Licensing fees for commercial solutions typically scale with the number of channels, the amount of content on those services being watermarked and/or the amount of activity from client applications (e.g. searches of an audio fingerprinting database). Finally, the lack of standards may result in high vendor switching costs.

Media synchronization has re-emerged as an active field of standardization during the last few years [5], several of which have been published in 2014. This article provides a comprehensive, but approachable, overview of recently published standards for media synchronization from the most relevant standardization bodies: IETF, ETSI, MPEG, DVB, HbbTV, and W3C.

Table 1 provides an overview of the standards for media synchronization discussed in this article. The following sections will provide further information about each of the standards, detail-

M. Oskar van Deventer and Hans Stokking are with the Netherlands Organisation for Applied Scientific Research TNO.

Matt Hammond is with BBC R&D.

Jean Le Feuvre is with Telecom ParisTech.

Pablo S. Cesar is with The National Research Institute for Mathematics and Computer Science CWI.

Body	Standard(s)	Year	Purpose	Timeline(s)	Correlation	Wall clock	Coordination
IETF	RFC 7272 [6]	2014	Social TV	RTP timestamp	N/A	Any (e.g. NTP)	RTCP extension
ETSI	TS 183 063 [7]	2011	IPTV, Social TV	RTP timestamp	RTCP extension	Same as IETF	RTCP extension
MPEG	TEMI [8]	2014	Timeline	MPEG TEMI	N/A	NTP / PTP	N/A
DVB	TS 103 286-2 [9]	2014	Companion screen	Various	CSS-MRS protocol	CSS-WC protocol	CSS-TS and CSS-TE protocols
HbbTV	TS 102 796 v1.3.1 [10]	2015	Hybrid TV, companion screen	Same as DVB	API	Same as DVB	DVB CSS-TS, API
W3C	SMIL [11]	2008	Multimedia presentations	Interactive	Semantic	N/A	SMIL scheduler

Table 1. Standards for media synchronization.

ing their main features, followed by a discussion about the current status of the standardization effort, and offering some insights on future developments.

Synchronizing RTP Streams: IETF RFC 7272

The IETF has standardized the use of the Real Time Control Protocol (RTCP) for the generic synchronization of Real Time Protocol (RTP) streams at different devices, e.g. social TV. RFC 7272 [6], published in 2014, defines the interaction between a synchronization client (SC) and a media synchronization application server (MSAS). The SC reports synchronization status information to the MSAS. The MSAS receives status information from multiple receivers in a single synchronization group. It can thus calculate a common playout time for all receivers. Based on this, the MSAS sends synchronization settings instructions to the SCs. A common clock that is synchronized across all receivers is assumed. No particular synchronization method is mandated, although several are suggested, including the Network Time Protocol (NTP).

Synchronization status information is defined in an RTCP extended report block, and contains, among others, the following information:

- A synchronization group identifier that distinguishes different groups of SCs that synchronize RTP streams.
- The packet-received RTP timestamp, identifying the RTP packet for which the timing is reported.
- The NTP time that a packet arrived at the input of the device.
- Optionally, the time that the contents of the packet were presented to the user.

The synchronization settings instructions are sent in a similar RTCP extended report. The main difference is that it contains a 64-bit packet-presented timestamp, allowing accurate control of presentation times. This is, for example, required in use cases such as audio beam forming or video wall displays.

Further, RFC 7272 specifies a Session Description Protocol (SDP) parameter that enables RTP entities to advertise their capability in various session control protocols. Furthermore, IETF has specified SDP parameters for negotiation of clock synchronization capabilities in RFC 7273.

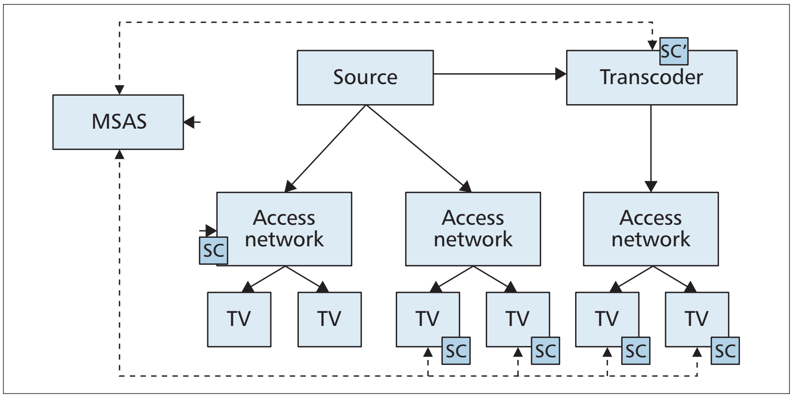

Figure 1. ETSI TS 183 063 functional entities for IDMS.

Social TV: ETSI TS 183 063

ETSI TISPAN specifies inter-device media synchronization as part of their Internet Protocol Television (IPTV) release 3 specifications, first published in 2011 [7]. The main use is social TV, in which various users can enjoy remote shared TV experiences, and which is spoiled by major path-delay differences between systems. The ETSI specification uses and expands upon RFC 7272.

A difference from RFC 7272 is that in the ETSI specification, the SC can be included in a TV, but it can also be part of an access network (see Fig. 1). This enables synchronization of large groups of legacy TVs by buffering at one node in the network. This requires that the delays from this node to those receivers are similar for all receivers, and that the SC compensates for those delays. Also, the MSAS can be a separate entity in the network, but it can also be contained in a TV. This enables a peer-to-peer style of synchronization.

To support this separation between function and element, ETSI specifies the session setup procedures. ETSI uses the Session Initiation Protocol (SIP) protocol for setting up media sessions, and uses the SDP protocol for specifying the synchronization part. In the SDP description exchanged between the TV and the IPTV system, the address of the MSAS is indicated. This can be a function in the network, or a function in another terminal.

An extension by ETSI to RFC 7272 is to support synchronization of the same content in different formats. In such a case, the RTP time-

The backbone of the media broadcasting industry is the MPEG-2 transport stream format, used throughout the broadcast chain. The MPEG-2 TS format specifies how video and audio or other media packets from one or several TV channels, called programs in MPEG-TS terminology, are scheduled within a continuous stream of bytes.

stamps reported by receivers are for different timelines. To support this situation, ETSI has defined a Synchronization Client *prime* (SC′), to be contained in a stream modifying entities such as a transcoder (see also Fig. 1). The SC' reports the mapping of timestamps in the incoming stream with those in the outgoing stream. This allows for conversion between those timestamps by the MSAS, thereby supporting SCs receiving original streams and SCs receiving modified streams that may be located in different networks.

External Media Items for Broadcast: MPEG TEMI

The backbone of the media broadcasting industry is the MPEG-2 transport stream (TS) format, used throughout the broadcast chain. The MPEG-2 TS format specifies how video and audio or other media packets from one or several TV channels, called programs in MPEG-TS terminology, are scheduled within a continuous stream of bytes. One challenge of the standard is to ensure that the target receiver behaves properly and consumes input data in time, avoiding buffer overflow, and to allow building systems with a limited amount of memory. This is achieved by a complex set of constraints on each bit in the byte flow, governed by the program clock reference (PCR) signal sent in the byte stream for each program. Decoding and presentation timestamps reference this clock. This forms an intrinsic timeline for the program, and is typically lost during transcoding or transmuxing of the source. To overcome this drawback, the MPEG group has defined an extension to MPEG-2 TS allowing for the carriage of extrinsic media clocks, along with other features, under the name 'TEMI' (Timing and External Media Information).

Earlier standards define such a mechanism using a dedicated elementary stream in the multiplex for the transport of an extrinsic clock. However, the method, while elegant, can be quite costly; at least one TS packet (188 bytes) has to be used to send the extrinsic clock of a program. When the clock has to be sent with each video frame for frame-accurate synchronization, this adds up to 75 kb/s for a 50 Hz video signal, more than the bandwidth used by some audio streams. TEMI addresses this issue by defining standard signaling, which is not affected by changes to PCR and that can be inserted before the start of a video or audio frame in the same TS packet. This design makes it possible to reduce the above signaling to 4 kb/s.

TEMI provides different ways of defining the extrinsic timeline:

- Sending a media time/media timescale pair, which can be compared with the presentation time of other media. Typically, this is compared with the composition time of an ISOBMFF track, or with the current time in an MPD period of an MPEG-DASH session.
- Sending an NTP or PTP (Precision Time Protocol) timestamp that matches timestamps associated with packets of the other media, for example interpolated NTP time of an RTP packet.
- Sending a time code of the media frame to be matched with a time code embedded in the other media, for example embedded as a track in ISOBMFF files or as an extension header in RTP packets.

TEMI also provides tools to signal the uniform resource location (URL) of one or several additional content items to be played synchronously with the broadcast, along with their MIME (Multi-Purpose Internet Mail Extensions) types. These items are assigned a timeline identifier, associated with each TEMI timing information. This allows sending URLs of associated services at a much lower frequency than timing information. Finally, TEMI provides a way of announcing when additional media content will become active by sending countdown signals for a given timeline identifier. TEMI content may also be marked as splicing points, indicating that the previous non-splicing content will resume at the end of the splice. This helps receivers to optimize their resources.

Figure 2 shows how TEMI may be used to signal and synchronize 3D or 4K enhancements (*URL#1*) and other additional contents such as subtitles or alternate audio (*URL#2*) to an existing broadcast signal, and signal upcoming splicing content (*URL#3*), typically for ad insertion purposes.

As a further continuation of TEMI activity, MPEG is investigating unified signaling of the different timelines defined in its various system layers (MPEG-4, ISO Base Media File Format, MPEG-DASH), along with specifying the missing tools enabling hybrid delivery of media content, such as signaling of coding dependencies between different containers.

Companion Devices: DVB CSS

The group Digital Video Broadcast – Companion Screens and streams (DVB-CSS) has developed a standard [9] to synchronize a media stream on a companion device with a media stream on a television set. The DVB-CSS architecture (Fig. 3) has one TV device and one or more companion screen applications (CSAs) running on companion screen devices that are connected via a home network, typically WiFi. Both TV and CSA independently receive media streams from the broadcaster (not shown). The presentation of the media streams is synchronized by using a set of new protocols and a new material resolution service.

A typical synchronization scenario is as follows. The user tunes their TV to a broadcast service. The TV receives the service, which includes a broadcast stream and metadata for media synchronization. The user pairs their companion device with the TV and starts a CSA. The TV provides the CSA with content identification and other information (CSS-CII protocol), which includes the service endpoints for the other protocols. The CSA queries the material resolution server (CSS-MRS protocol) and obtains material information that describes the structure of the broadcast (composition of materials and sub-materials such as programs, sections within programs, and adverts). It also describes the relationship between this structure and timelines. DVB-CSS supports several types of timelines,

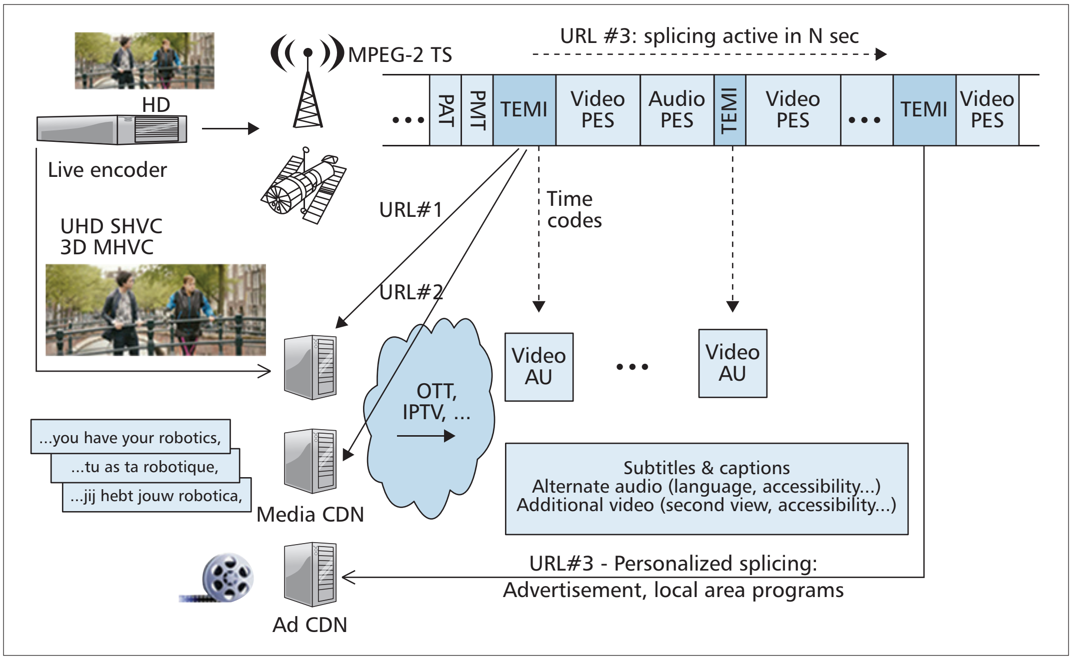

Figure 2. TEMI usage in hybrid broadcast.

HbbTV is an industry forum that specifies a HTML+JavaScript application programming interface (API) for browser-based applications on TVs, launched in 22 countries as of March 2015. The new HbbTV 2.0 specification [10] includes features for media synchronization, which are a profile of DVB CSS.

including MPEG transport stream presentation timestamp, ISOBMFF composition time and time relative to the start of a period in an MPEG DASH presentation. It also supports the use of TEMI as a timeline, but not as a means of specifying the external media stream to be played by the companion screen.

This combination of information from the TV and a server enables the CSA to determine which streams it should present and how its timeline correlates to that of the media being presented by the TV. However, the CSA manages its own behavior and is not directly controlled by the TV.

In parallel, the CSA synchronizes its wall clock with the TV (CSS-WC protocol). When the user starts a selected media stream on its companion screen, the CSA synchronizes the stream's timeline with the timeline of the broadcast stream on the TV (CSS-TS protocol described below). The CSA can also subscribe to trigger events (CSS-TE protocol) that are received by the TV from the broadcaster as part of the signaling within the broadcast stream.

The Wall Clock Synchronization protocol (CSS-WC) is a request-response UDP-based protocol that enables the client (CSA) to estimate a clock at a server (the TV) and measure and compensate for network round-trip delay. The protocol design is similar to the client/server mode of NTP but significantly simplified. Although many devices implement NTP to set their system-wide clocks, a CSA running on a device cannot always check if an NTP client process is functioning or query the accuracy of clock synchronization. Media synchronization also does not require the shared clock to be with reference to absolute real world time and can therefore avoid complexities such as leap-seconds. Frame accurate media synchronization requires accuracy of the order of milliseconds, and the chances of achieving this are improved if the protocol operates directly between the TV and CSA instead of via a hierarchy of intermediate servers on more distant network segments.

The Timeline Synchronization protocol (CSS-TS) is a websocket-based protocol that carries the timing information needed for coordination between the CSA and TV. Messages conveyed by this protocol describe the relationship between wall clock time and timeline position. This enables the CSA to accurately estimate the current TV timeline position despite possible network transmission delays. Timeline positions reported by the TV are expected to take account of any delays between the point at which it samples the timeline position in its media pipeline and the display of the media. Similarly, if a set-top box and an HDMI-connected display is used, then the STB is expected to make a best-effort to compensate for the play-out delay of the display. HDMI signaling may be used for this purpose.

Hybrid TV: HbbTV 2.0

HbbTV is an industry forum that specifies a HTML+JavaScript application programming interface (API) for browser-based applications on TVs, launched in 22 countries as of March 2015. The new HbbTV 2.0 specification [10] includes features for media synchronization, which are a profile of DVB CSS. In an HbbTV 2.0 TV, media synchronization is possible between the TV and a CSA or another HbbTV 2.0 TV acting in the role of a CSA. It is implemented as a profile of the DVB-CSS specification, and it is only activated when an interactive application running on the TV explicitly requests it. The DVB-defined CSS-CII, CSS-WC, and CSS-TS protocols are used in HbbTV 2.0, but HbbTV 2.0 TVs are not required to implement CSS-TE.

Media synchronization functionality between TV and CSA is available for most media types that the TV can be playing, including both broadcast and streamed broadband content. Required support for media synchronization between mul-

Synchronized multimedia integration language and nested context language are the most relevant examples of declarative and structured rich media formats. SMIL has been standardized by the World Wide Web Consortium, while NCL is the multimedia presentation standard for IPTV selected by ITU (ITU H.761).

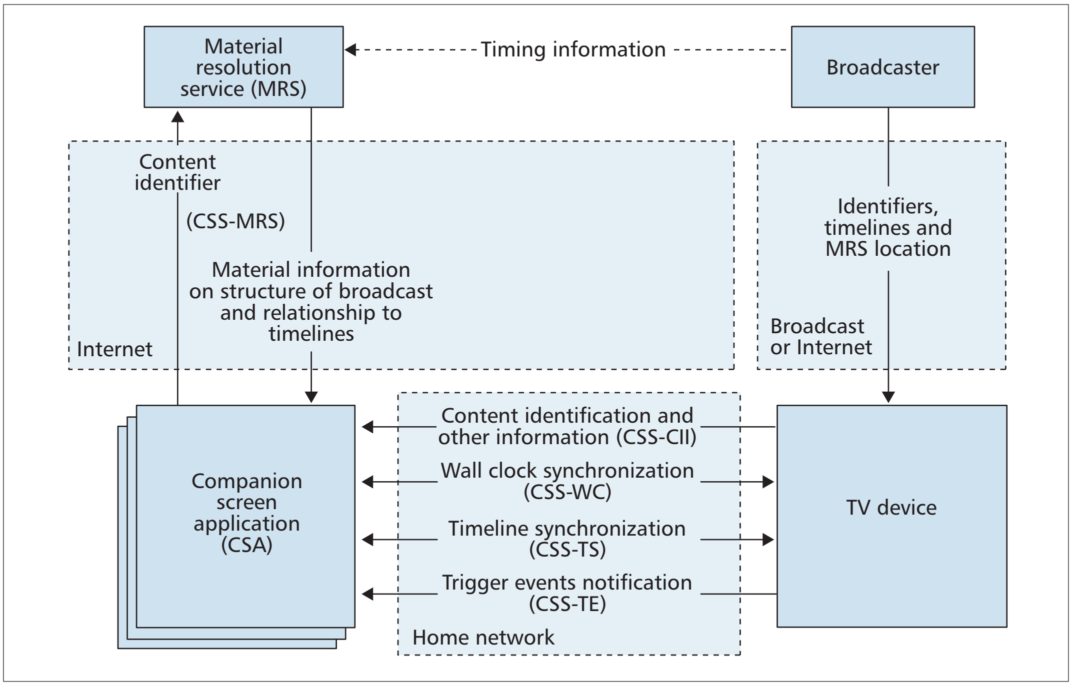

Figure 3. DVB-CSS architecture [9].

tiple streams within the TV is limited to combinations where one stream (possibly together with a synchronizing subtitle data stream) is delivered by broadcast and another stream by broadband.

HbbTV 2.0 specifies a single API that can be used for both single-TV multi-stream synchronization and inter-device synchronization. For the latter, an HbbTV terminal can act as both "master" and "slave," enabling streams on two TVs to be synchronized with each other. The API controls the life cycle of the MediaSynchoniser JavaScript object. This object is initialized by the API and populated with media objects, corresponding to to-be-synchronized media streams. The API also has methods to enable and disable the inter-device protocols explained above.

A media synchronization buffer is optional in HbbTV 2.0. Even without a buffer in the TV, media synchronization may be possible. The broadcaster can preload media streams for the CSA (or a slave HbbTV terminal) on a CDN. The broadcaster could editorially delay the broadcast stream, although this is not typically done for live streams. If any of the media streams is MPEG DASH (HbbTV only supports this type of standards-based adaptive streaming), then it is mandatory to buffer the stream on a CDN. If the TV has a media-sync buffer, then it will be at least 30 MByte large. This is sufficient to reliably buffer at least 10 seconds of encoded high definition television (HDTV) content.

Multimedia Presentations: W3C SMIL and ITU NCL

Synchronized multimedia integration language (SMIL) and nested context language (NCL) are the most relevant examples of declarative and structured rich media formats. SMIL has been standardized by the World Wide Web Consortium (W3C) [11], while NCL is the multimedia presentation standard for IPTV selected by ITU (ITU H.761) [12]. They are both XML-based integration formats, and as such they do not directly define media objects. Instead, they define the temporal and spatial relationships between the different media objects, enabling media synchronization of distributed objects across heterogeneous devices. They both sit on top of other low-level transmission and delivery standards, which are in charge of executing the low-level synchronization primitives.

The core part of these standards is the scheduler. The scheduler is in charge of constructing a time graph of the presentation, based on the duration of the media items and on the temporal synchronization between them. Based on the time graph, media items composing the presentation become active or inactive at specific moments in time. The scheduler is dynamic, allowing for the description of adaptable presentations based on events (from the user, from the network, from third-party entities like a broadcaster). They can be used for a variety of use cases from social television applications (including secondary screen support) to video conferencing services, to late-binding mashup videos [13].

NCL and SMIL have a strict separation between the document's content and structure, and it provides non-invasive control of presentation timing, linking, and layout. In NCL, authors can declaratively describe the temporal behavior of a multimedia presentation using connectors and links. SMIL acts as a container format in which spatial, temporal, linking, and interactive primitives can be used to position, schedule, and control a wide assortment of multimedia presentations. Both languages also allow for some form of procedural control. The biggest difference between the two languages is that while SMIL provides high-level constructs defining a restricted set of temporal relationships, NCL allows an author to create a set of custom relationships from a toolkit of language primitives as objects.

Both languages incorporate the recurrent aspects from a multimedia presentation [14]:

- Media items: defining what to render (video, images, text, and sometimes 3D objects). For selectivity purposes, the document model might also provide mechanisms for rendering one of multiple alternative assets.
- Style: defining how to render media, including multimedia styling options and digital effects such as zooming within an image.
- Spatial composition: defining where to render media in order to provide a meaningful and aesthetically attractive presentation.
- Temporal composition: defining when to render media including the start time and duration of media items, and also synchronization constraints between the items.
- User interaction: defining how to influence the presentation.

Conclusion

Current standardization efforts target specific use cases, such as social TV, hybrid broadcast/broadband services, and companion screens, which require media synchronization. This article provides an overview of them, highlighting the current industry push for such new services, both at the IP media stream level (IETF RTCP, ETSI TISPAN) and the MPEG-2 transport stream level (DVB CSS, MPEG TEMI). It also includes more fundamental standards (W3C SMIL and ITU NCL) that can serve as models for future more general synchronization primitives.

Successful standardization efforts are key for industrial partners. Vendors in HbbTV have already committed to implement at least the mandatory aspects of HbbTV 2.0 (including the profile of DVB-CSS) in their new TV products, with the expectation of seeing compliant products by 2017. Meanwhile broadcasters, including the BBC, are already exploring the services that media synchronization will enable [15].

References

[1] D. Geerts *et al.*, "Are We in Sync? Synchronization Requirements for Watching Online Video Together," *ACM Conf. Human Factors in Computing Systems, Vancouver (ACM-SIGCHI)*, May 2011.

[2] ITU-R BT.1359-1, "Relative Timing of Sound and Vision for Broadcasting," 1998.

[3] W.J. Kooij *et al.*, "Playout Delay of TV Signals: Measurement System Design, Validation and Results," *ACM Int'l. Conf. Interactive Experiences for TV and Online Video (ACM-TVX)*, June 2014, pp. 23–30.

[4] BBC R&D white paper 101, "EBU Tests of Commercial Audio Watermarking Systems," Dec. 2004.

[5] M. Montagud *et al.*, "Inter-Destination Multimedia Synchronization: Schemes, Use Cases and Standardization," *Multimedia Systems*, vol. 18, no. 6, 2012, pp. 459–82.

[6] IETF Inter-Destination Media Synchronization (IDMS) Using the RTP Control Protocol (RTCP), IETF RFC 7272, June 2014.

[7] ETSI IMS-based IPTV Stage 3 Specification, ETSI TS 183 063, Mar. 2011.

[8] MPEG TEMI, ISO/IEC 13818-1:201x/FDAM 2 Delivery of Timeline for External Data, Dec. 2014.

[9] DVB Companion Screens and Streams; Part 2: Content Identification and Media Synchronisation, DVB Document A167-2, July 2014, also published as ETSI TS 103 286-2.

[10] HbbTV Hybrid Broadcast Broadband TV version 2, Feb. 2015, available at hbbtv.org, to be published as ETSI TS 102 796 v1.3.1.

[11] W3C Synchronized Multimedia Integration Language, W3C SMIL 3.0, Dec. 2008.

[12] ITU-T H.761, "Nested Context Language (NCL) and Ginga-NCL for IPTV Services," 2009.

[13] R.L. Guimaraes *et al.*, "Creating Personalized Memories from Social Events: Community-Based Support for Multi-Camera Recordings of School Concerts," *ACM Multimedia Conf. (ACM MM 2011)*, Scottsdale (AZ), USA, Nov. 28–Dec. 1, 2011, pp. 303–12.

[14] J. Jansen, P. Cesar, and D. Bulterman, "A Model for Editing Operations on Active, Temporal Multimedia Documents," *ACM Symp. Document Engineering (ACM DocEng 2010)*, Manchester, UK, Sept. 21–24, 2010, pp. 87–96.

[15] S. Jolly and M. Evans, BBC R&D white paper 242 (Feb. 2013): "Improving the Experience of Media in the Connected Home with a New Approach to Inter-Device Communication."

Biographies

M. Oskar van Deventer (oskar.vandeventer@tno.nl) received M.Sc. and Ph.D. degrees in electrical engineering from the Eindhoven Technical University in 1987 and 1994, respectively. He became a Member (M) of IEEE in 2009. He has worked at KPN Research since 1987 and at TNO since 2003, where he is a senior scientist media networking. His current focus is on the realization of international standards for media networking, including DVB, HbbTV, ETSI, IETF and OIPF. He is the author of one book, and (co)author of more than 100 publications, 60 patent applications, and 600 standardization contributions.

Hans Stokking received the M.Sc. degree in systems engineering, policy analysis & management from the Delft Technical University in 1999. He has worked at KPN Research since 1999 and at TNO since 2003, where he is a senior scientist media networking. His work focuses on multi-domain networking and delivery of IPTV services, with a special research interest in media synchronization. He is the organizer of the Media Synchronization Workshop series and TPC member for various conferences.

Matt Hammond received an M.Eng. degree in computer systems and software engineering from the University of York in the United Kingdom in 2001. From 2001 to 2011 he worked as a research engineer in the Research and Development Department of the BBC in London in the United Kingdom, and as a senior research engineer since 2011. His current work focus is on companion device interaction with televisions and broadcast services. He is an active contributor to the DVB Companion Screens and Streams specification and to HbbTV in the areas of synchronization and other companion screen functionality.

Jean Le Feuvre received his M.Sc. degree in telecommunications in 1999 from TELECOM Bretagne. He is the co-founder of Avipix, an MPEG-4 consulting and software provider company. He has been involved in MPEG standardization since 2000. He joined TELECOM ParisTech in 2005 as associate professor in the Signal Processing and Image Department. He is an expert in multimedia delivery and interactive video systems. He is the project leader and maintainer of GPAC, a multimedia framework based on standards technologies (MPEG, IETF, W3C, and Web3D). He is the editor of several ISO/IEC standards, and the co-author of several books, more than 30 publications and patents applications, and 300 standardization contributions.

Pablo Cesar received his Ph.D. from Helsinki University of Technology in 2005. He leads the Distributed and Interactive Systems Group at Centrum Wiskunde & Informatica (CWI). He has (co)authored more than 70 articles (conference papers and journal articles) about multimedia systems and infrastructures, social media sharing, interactive media, multimedia content modeling, and user interaction. He is involved in standardization activities (e.g. W3C, MPEG, ITU) and has been active in a number of European projects. He is co-editor of the book *Social Interactive Television: Immersive Shared Experiences and Perspectives*, and has given tutorials about multimedia systems in prestigious conferences such as ACM Multimedia and the WWW conference.

Successful standardization efforts are key for industrial partners. Vendors in HbbTV have already committed to implement at least the mandatory aspects of HbbTV 2.0 (including the profile of DVB-CSS) in their new TV products, with the expectation of seeing compliant products by 2017.

COMMUNICATIONS STANDARDS

A Supplement to IEEE Communications Magazine

MARCH 2016

SUPPLEMENT EDITOR-IN-CHIEF
GLENN PARSONS
SUPPLEMENT ASSOCIATE EDITOR-IN-CHIEF
HUUB VAN HELVOORT
MANAGING EDITOR
JACK HOWELL
STANDARDS NEWS CONTRIBUTORS
NIGEL DAVIS • MICHAEL BRENNER • HING-KAM LAM • SCOTT MANSFIELD
KENNETH DILBECK • STEVEN WRIGHT