

Plataforma extremo-a-extremo compatible con el estándar HbbTV 2.0 para la TV híbrida y multi-dispositivo

Dani Marfil, Fernando Boronat, Mario Montagud, Pau Salvador.

Departamento de Comunicaciones, Immersive Interactive Media (IIM) R&D Group,

Universitat Politècnica de València - Campus de Gandia

C/ Paranimf, 1 46730 Grau de Gandia (Valencia).

{damarre@dcom., fboronat@dcom., mamontor@, pasallla@epsgr.}upv.es

Resumen- En este artículo se presenta una plataforma extremo-a-extremo para la generación, entrega y sincronización de contenidos híbridos (*broadcast* y *broadband*) relacionados, tanto en un mismo como en múltiples dispositivos (ej. escenario multi-pantalla). La plataforma es compatible con la versión más reciente del estándar *Hybrid Broadcast Broadband TV* (2.0.1) y, además, incluye soluciones eficientes para aspectos clave que no están especificados en dicho estándar, pero que son necesarios para la implantación satisfactoria de servicios de TV híbridos y multi-dispositivo, como: mecanismos de señalización para descubrir, asociar y describir los contenidos (híbridos) adicionales relacionados, mecanismos de interacción y negociación entre los dispositivos disponibles y soluciones de sincronización multimedia adaptativa (incluyendo protocolos, algoritmos y técnicas de ajuste) para alinear en el tiempo, de manera precisa, la presentación de los contenidos disponibles. La plataforma se ha evaluado objetiva y subjetivamente para el caso de uso de la TV multi-vista y multi-dispositivo, obteniendo resultados satisfactorios y prometedores.

Palabras Clave- HbbTV, Sincronización Híbrida, IDES, Sincronización Multimedia, Escenario Multi-Dispositivo, TV Multi-Vista, DVB, DASH, QoE

I. INTRODUCCIÓN

En la actualidad, una gran variedad de tecnologías de distribución de contenidos, de dispositivos de consumo, así como de contenidos multimedia, está al alcance de los consumidores. En cuanto a las tecnologías de distribución, los contenidos multimedia se pueden enviar bien a través de redes de radiodifusión, en adelante, redes *broadcast*, como, por ejemplo, DVB (*Digital Video Broadcasting*), o bien a

través de redes IP, en adelante, redes *broadband*, como, por ejemplo, Internet.

Por una parte, las tecnologías *broadcast* pueden proporcionar el mismo contenido multimedia a un gran número de consumidores. Por otra parte, las tecnologías *broadband* pueden proporcionar servicios bidireccionales, interactivos y adaptativos, en base a las preferencias y recursos de los consumidores. Sin embargo, suelen presentar un rendimiento más pobre en términos de escalabilidad, estabilidad y latencia, en comparación con las tecnologías *broadcast*. En este contexto, el contenido puede ser entregado empleando diversas variantes de descarga y de transmisión (*streaming*) de contenidos, siendo esta última opción la que ha ganado mayor popularidad en los últimos años. Respecto a tecnologías o servicios de *streaming*, pueden distinguirse dos alternativas principales: gestionados y no gestionados [1]. Los servicios gestionados operan típicamente en entornos controlados como, por ejemplo, IPTV (*Internet Protocol TV*), y se basan principalmente en *push-based streaming*, empleando los protocolos RTP/RTCP (*Real-time Transport Protocol / RTP Control Protocol*) [2]. Dichos protocolos son especialmente apropiados para servicios interactivos, en los que el retardo no debe superar ciertos umbrales. Los servicios no gestionados pueden operar en entornos de área extensa y están típicamente basados en *pull-based streaming*, utilizando HTTP (*HyperText Transfer Protocol*) como mecanismo de descarga adaptativa, siendo típicamente conocidos como HAS (*HTTP-based Adaptive Streaming*). Las principales ventajas que ofrecen son la adaptabilidad, escalabilidad, fiabilidad, ubicuidad y

eficiencia en el coste. En este contexto, diferentes empresas y organismos de estandarización han especificado su propia solución HAS, entre las cuales destacan: *HTTP Live Streaming* (HLS) de Apple; *MPEG Dynamic Adaptive Streaming over HTTP* (DASH) de ISO/IEC y MPEG [3]; y *Smooth Streaming* de Microsoft (MSS). Las soluciones HAS están siendo mejoradas continuamente y ampliamente adoptadas para la distribución de contenidos multimedia en redes *broadband*. Como prueba de evidencia, DASH ha sido adoptado por el estándar *Hybrid Broadcast Broadband TV* (HbbTV) [4] y está siendo adoptado por servicios y plataformas muy populares.

De algún modo, las tecnologías *broadcast* y *broadband* se han convertido en rivales en el mercado actual de distribución y consumo de contenidos multimedia. Sin embargo, la inter-operabilidad, coordinación y convergencia entre ambas tecnologías, en combinación con la amplia disponibilidad de dispositivos multi-conectados, abre la puerta a nuevas posibilidades de innovación y oferta de servicios multimedia enriquecidos. Algunos ejemplos de estos servicios son: el consumo en paralelo de diferentes flujos de vídeo (ej. el modo *Picture-In-Picture* -PiP- o el modo mosaico), provisión de escalabilidad espacial, temporal o de color [5], *tiled-streaming* (ej. distribución de vídeo UHD -*Ultra High Definition*-, donde diferentes áreas espaciales del mismo vídeo se entregan a través de diferentes flujos) [6], selección personalizada de flujos de audio (ej. sustitución del audio del contenido *broadcast* por uno proveniente de una emisora de radio *on-line*), etc.

Debido al potencial que ofrece la coordinación y convergencia entre las tecnologías *broadcast* y *broadband*, se persigue disponer de un ecosistema multimedia híbrido en el que ambas tecnologías se complementen y aumenten su valor unidas. Una prueba de evidencia es el estándar HbbTV [4], que proporciona mecanismos para armonizar la entrega y consumo interactivo de contenidos *broadcast* y *broadband* a través de TV conectadas y dispositivos secundarios con conectividad IP.

Sin embargo, la distribución y consumo de contenidos multimedia a través de tecnologías híbridas todavía se enfrenta a numerosos retos, como: la configuración y adaptación de dispositivos a esta reciente tecnología, cuya última especificación [4] aún no ha sido comercializada; la especificación e implementación de soluciones para la señalización y descubrimiento de los servicios multimedia híbridos disponibles y las aplicaciones relacionadas; el diseño e implementación de mecanismos de descubrimiento e interacción entre los dispositivos involucrados en la sesión multimedia; el diseño e implementación de soluciones de sincronización multimedia adaptativa y precisa para alinear en el tiempo el consumo de los contenidos híbridos, tanto en escenarios mono- como multi-dispositivo, etc.

Como principal contribución de este trabajo, se presenta una plataforma compatible con el estándar

HbbTV 2.0.1 que incluye soluciones eficientes para los retos anteriormente identificados. Además, la plataforma se ha evaluado, tanto objetiva como subjetivamente, para el caso de uso de TV multi-vista y multi-dispositivo que, tal y como se argumenta en las próximas secciones, es un caso de uso bastante relevante y con potencial comercial.

El artículo se estructura de la siguiente manera: en la sección 2, se revisan las soluciones de sincronización híbrida e IDES (*Inter Device Media Synchronization*) más relevantes. En la sección 3, se presenta la plataforma desarrollada, describiendo sus componentes y funcionalidades. En la sección 4, se presentan algunos resultados de las evaluaciones objetivas y subjetivas realizadas. Finalmente, en la sección 5, se presentan las conclusiones y algunas ideas para trabajo futuro.

II. ESTADO DEL ARTE

A lo largo de los años, se han propuesto diferentes soluciones para la sincronización de contenidos multimedia, utilizando distintas tecnologías de transmisión, en diferentes entornos de red y aplicaciones. Asimismo, numerosos estudios han analizado los avances con respecto a la sincronización multimedia. Una clasificación de soluciones de *sincronización inter-media* e *Inter-Destination Media Synchronization* (IDMS) puede encontrarse en [7], mientras que en [8] se proporciona un repaso histórico sobre la sincronización multimedia, teniendo en cuenta los avances tecnológicos, modelos teóricos y estudios sobre la percepción humana. El estudio en [9] proporciona una revisión de los estándares más recientes para la *sincronización inter-media, híbrida e IDMS*. Además, en [10] se analiza cómo las referencias de reloj y las marcas de tiempo se utilizan en diferentes estándares MPEG y DVB.

Esta sección se centra principalmente en la revisión de los trabajos más relevantes en el contexto de la sincronización híbrida. Además, también se describen brevemente las pruebas de concepto más relevantes implementadas para IDES. Las soluciones de sincronización basadas en técnicas propietarias (como *watermarking* o *fingerprinting*) no se consideran, debido a sus múltiples inconvenientes (baja precisión, sobrecarga, sensibilidad al ruido, escalabilidad baja...), tal y como se describe en [9], además de no ser apropiadas en el contexto de este trabajo.

En [11] se propone un algoritmo para sincronización *inter-media* entre flujos multimedia generados por el mismo proveedor de contenidos. Dicho algoritmo se basa en controlar la reproducción del audio y del vídeo teniendo en cuenta el valor de los campos PCR (*Program Clock Reference*) y PTS (*Presentation Timestamp*) asociados, incluidos en los flujos MPEG2-TS. El uso de los campos PCR y PTS para sincronizar diferentes flujos también se analiza en [12]. Sin embargo, en este caso los flujos son generados por distintas fuentes que comparten una fuente de reloj común (ej., utilizando *Network Time Protocol*, NTP)

para insertar valores PTS en los flujos MPEG2-TS de manera sincronizada.

En [13] también se aborda la transmisión híbrida *broadcast/broadband* de contenidos audiovisuales utilizando MPEG2-TS. En dicho trabajo, se identifica el hecho de que los codificadores *broadcast* y *broadband* no compartan sus relojes como una barrera para conseguir la sincronización híbrida. Como respuesta, se propone unificar la transmisión *broadcast* y *broadband* a través de la utilización del protocolo IP en ambas e insertando marcas de tiempo comunes.

En [14], el concepto de contenidos híbridos se analiza desde un punto de vista diferente: se investiga cómo segmentar un único flujo y programar la transmisión de los distintos segmentos a través de todas las redes disponibles, con tal de garantizar una reproducción apropiada en la parte del consumidor, con el mínimo número de interrupciones.

En [15] también se investiga sobre el consumo sincronizado de contenidos híbridos. En particular, dicho trabajo se centra en la sincronización de un flujo de contenido audiovisual con formato MPEG2-TS transmitido a través de un canal IPTV, empleando los protocolos RTP/RTCP, y un flujo de audio con formato MP3, transmitido vía *Internet Radio*, empleando también los protocolos RTP/RTCP. El objetivo de la sincronización se consigue asegurándose que todos los proveedores de contenido involucrados insertan marcas de tiempo provenientes de una fuente de reloj común, mediante el uso de NTP, y, posteriormente, ejecutando dos procesos principales: sincronización inicial y sincronización continua.

En [16] se propone una solución diferente para la sincronización híbrida. Se basa en el uso de relojes globales, no requiere la existencia de un canal de retorno y no requiere comunicación entre las redes y/o tecnologías utilizadas. Se investigan dos escenarios. Por un lado, la sincronización entre un flujo de audio transmitido vía *broadcast* FM y un flujo audiovisual MPEG2-TS transmitido vía *broadband* se consigue mediante el uso de un reloj UTC (*Universal Time Clock*) común e insertando marcas de tiempo de dicho reloj en las estructuras RDS (*Radio Data System*) del flujo FM y en la tabla *Time and Date Table* (TDT) del flujo MPEG2-TS. Por otro lado, también se consigue la sincronización entre un MPEG2-TS transmitido vía *broadcast* y otro vía *broadband*. En dicho caso, se logra sincronizar ambos flujos mediante la inserción de marcas de tiempo en las tablas TDT de cada uno de ellos.

En [18] se muestra que se puede conseguir sincronización a nivel de trama entre contenidos híbridos en un único dispositivo, empleando para ello flujos MPEG2-TS locales y un flujo externo MPEG-DASH. Para ello, utilizan la solución basada en PCR/PTS en combinación con una solución propuesta por la ETSI [19] para proporcionar líneas de tiempo absolutas en los flujos MPEG2-TS. Esta última solución permite insertar líneas de tiempo extrínsecas y absolutas dentro de un flujo MPEG2-TS, consiguiendo

de esta forma un reloj común para diferentes flujos. Así, se superan los problemas que surgen cuando únicamente se confía en la solución basada en PCR/PTS (como en [11] y [12]). En dicho trabajo, se implementa un banco de pruebas utilizando el *framework* multimedia GStreamer [20] (también empleado en la plataforma propuesta).

El trabajo en [21] es una evolución del trabajo en [16], y presenta una solución para la transmisión de contenidos híbridos (DVB y DASH), basada en la solución propuesta por la ETSI [19]. En dicho trabajo se extiende la plataforma GPAC [22] (también utilizada en la plataforma presentada) para: un mejor soporte de DASH; poder identificar la localización del contenido *broadband* (DASH) adicional; así como incluir una descripción del contenido que está siendo transmitido. Asimismo, el trabajo en [5], enumera y describe casos de uso relevantes relacionados con la transmisión de contenidos híbridos y presenta un banco de pruebas, también basado en GPAC, para implementar uno de ellos. En dicho trabajo se utiliza un mecanismo más reciente para conseguir sincronización híbrida, bajo el nombre de TEMI (*Timing External Media Information*) [17], propuesto por MPEG y DVB como una mejora (*amendment*) a la especificación ISO/IEC 13818-1. Este mecanismo cuenta con las siguientes características: 1) inserción de una línea temporal extrínseca, absoluta y estable en el flujo MPEG2-TS; 2) inserción de la localización (vía URL) del contenido *broadband* relacionado en el flujo MPEG2-TS; 3) notificación del momento de disponibilidad del contenido *broadband*. Los dos primeros tipos de metadatos se conocen comúnmente como *TEMI timeline* y *location descriptor*, respectivamente. En particular, en [5] se añade soporte para la inserción y extracción de marcas TEMI en el multiplexor y demultiplexor MPEG2-TS, así como para la codificación *High Efficiency Video Coding* (HEVC) y su extensión escalable multi-nivel o multi-capa (*Layered-HEVC*). Mediante este esquema de codificación y transmisión híbrida, la capa base HEVC se envía vía *broadcast* (DVB) y las capas de mejora vía *broadband* (DASH) siendo recibidas bajo demanda por los consumidores.

Nuestro trabajo también se basa en el uso de TEMI para conseguir sincronización híbrida, por las ventajas que ofrece sobre el mecanismo propuesto por la ETSI [19], tal y como se describe en [9], y por ser el mecanismo adoptado por el estándar HbbTV.

Todos los trabajos mencionados hasta este punto han abordado la sincronización híbrida en un único dispositivo, pero la sincronización IDES también se está convirtiendo en un tema de especial relevancia, debido a la masiva proliferación y uso de dispositivos secundarios, tales como *smartphones* y *tablets*. En este contexto, el trabajo en [23] implementa el mecanismo propuesto por la ETSI basado en líneas temporales absolutas [19] como una solución para proporcionar sincronización híbrida en escenarios TV multi-pantalla, utilizando dispositivos secundarios. Las pruebas de

evaluación muestran que la precisión obtenida en la sincronización entre un flujo de audio *multicast* y un flujo de vídeo DVB-T es suficiente para conseguir *lip-sync* (precisión del orden de 80 ms). En [24], se combinan un conjunto de componentes tecnológicos (mecanismo definido por la ETSI [19], protocolo IDES, mecanismos de descubrimiento y solución para *tiled-streaming*) con el objetivo de proporcionar experiencias multi-pantalla inmersivas. En concreto, se desarrollan aplicaciones HbbTV (previas a la especificación de la versión 2.0) para sincronizar contenidos DVB-T en un terminal híbrido con contenidos DASH (*tiled-streaming*) en dispositivos secundarios. Sin embargo, el contenido *broadcast* se “simula” en dicho trabajo con contenido local, sin realizar transmisión y recepción DVB-T.

III. PLATAFORMA COMPATIBLE CON HBBTV 2.0

La plataforma extremo-a-extremo para la distribución y consumo sincronizado de contenidos multimedia híbridos desarrollada en este trabajo se divide en dos partes: la de proveedor(es) de contenidos y la del usuario/consumidor final. La Fig. 1 proporciona una visión general de la plataforma.

La plataforma es válida para entornos mono- y multi-dispositivo, en los que existe un dispositivo principal, en adelante MS (*Main Screen*) y uno (o varios) dispositivos secundarios, en adelante CS (*Companion Screen*), integrados en el mismo dispositivo que el MS o bien en dispositivos secundarios (ver Fig. 1). El MS consume contenidos MPEG2-TS recibidos vía *broadcast* (DVB), sobre un evento específico (p.ej., un concierto), en una TV principal conectada (terminal híbrido). Este flujo *broadcast*, además, contiene información insertada (metadatos) que permite que los CS activos puedan consumir contenidos multimedia adicionales relacionados (p.ej., vídeos relacionados con el contenido *broadcast*) distribuidos vía *broadband* (DASH).

La plataforma se ha desarrollado empleando, en su mayor parte, el *framework* GStreamer [20], tanto en dispositivos basados en Linux como en Android.

A. Parte de Proveedor(es) de Contenidos

En la parte del proveedor de contenidos (zona izquierda de la Fig. 1) la plataforma incluye las funcionalidades necesarias para la codificación, preparación, segmentación, almacenamiento, modulación y transmisión de los contenidos multimedia. Asimismo, esta parte también incluye funcionalidades para la generación, inserción y almacenamiento de información o metadatos (ej. líneas de tiempo absolutas, descripción y localización de los contenidos multimedia disponibles, etc.) necesarios para conseguir sincronización híbrida.

La plataforma permite la entrega de contenidos multimedia vía *broadcast*, empleando DVB-T, y vía *broadband*, empleando DASH, que es la tecnología adoptada por HbbTV para la distribución de contenidos *broadband*. Además, también soporta otras tecnologías como HLS o *Real Time Streaming Protocol* (RTSP) + RTP/RTCP.

Con respecto a *broadcast*, la plataforma permite la configuración de la codificación de los contenidos, la inserción de los descriptores TEMI (empleando herramientas de GPAC [22]) y la modulación y transmisión de los contenidos generados, utilizando una tarjeta PCI Dektec DTA-2111 y su software asociado *StreamXpress*. La tasa de inserción de los descriptores TEMI es configurable (aunque típicamente se añade el *TEMI timeline descriptor* una vez por cada trama).

Con respecto a *broadband*, la plataforma soporta todos los requisitos necesarios para ofrecer contenidos DASH, empleando una herramienta desarrollada por nuestro grupo en un trabajo previo [25]. En concreto dichos requisitos son: 1) codificación en diferentes calidades (*bitrates*, resoluciones, patrones de GoP - *Group of Pictures*-...); 2) segmentación, con un tamaño de segmentos (*chunks*) configurables; 3) generación del *Media Presentation Description* (MPD); y 4) almacenamiento del contenido DASH y del MPD en un servidor web convencional. Además, se han realizado procesos similares para añadir soporte para HLS, mientras que también se ha desarrollado un servidor RTSP + RTP/RTCP, empleando para ello librerías y plugins disponibles en el *framework* GStreamer [20].

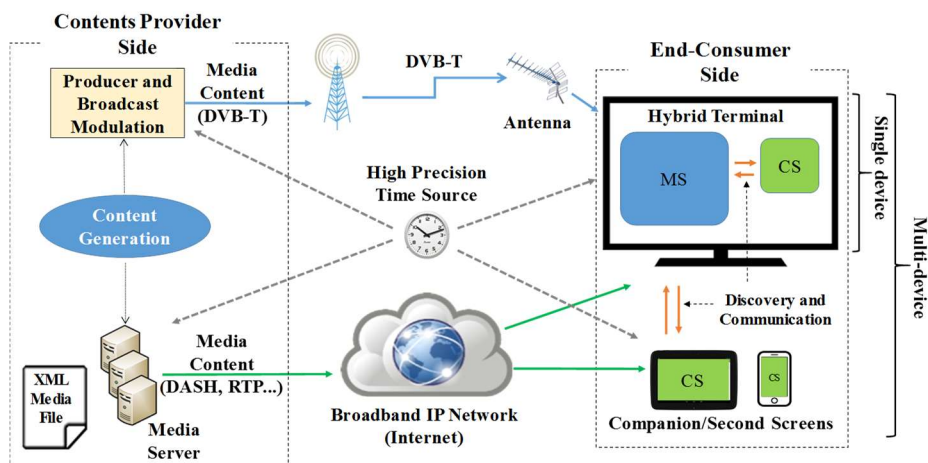


Fig. 1 Visión general de la plataforma desarrollada y del escenario evaluado

Aparte de la generación y preparación de los contenidos multimedia híbridos, así como de la inserción de descriptores TEMI, también es imprescindible la inclusión de mecanismos adicionales para indicar la disponibilidad de dichos contenidos multimedia relacionados y para proporcionar los metadatos relevantes sobre los mismos (ej. su relación con el contenido *broadcast*, la naturaleza del contenido, las URL para acceder a los mismos...). De esta forma los receptores pueden conocer su existencia y disponer de la información necesaria para decidir si quieren o no reproducirlos y, en caso afirmativo, preparar los reproductores adecuados. Para el descubrimiento de los contenidos y la notificación de su existencia, se emplea el *TEMI location descriptor*. Como mecanismo de enlazado y descripción de contenido, se emplea un fichero XML, con los nodos y atributos necesarios. Este fichero XML se almacena en un servidor web (ej. el mismo servidor que almacena el contenido *broadband*), al cual se accede vía HTTP utilizando la URL insertada en el *TEMI location descriptor*. Dependiendo de la naturaleza de los contenidos híbridos disponibles, se incluyen metadatos relevantes específicos, como por ejemplo la tecnología de distribución (DASH, HLS, RTP...), los tipos de contenido (audio, video, web), su formato de encapsulación (MP4, MPEG2-TS...), su codificación (H264, MP3...) y otros metadatos relevantes (ej. una breve descripción del contenido, idioma del audio, etc.). Además, el fichero XML también incluye otro tipo de información relevante, como la fecha de su última modificación, así como la tecnología utilizada para obtener una fuente de reloj global (ej. *Precision Time Protocol -PTP-*, NTP...), junto con la URL del servidor utilizado. Esta fuente de reloj global se utilizará para proporcionar una noción coherente del tiempo en la sesión multimedia, ya sea en la inserción de marcas temporales TEMI o en su interpretación/intercambio en la parte del consumidor final, posibilitando así el objetivo de la sincronización híbrida. Este fichero XML es descargado y analizado por el terminal híbrido en la parte del usuario final.

B. Parte del Usuario/Consumidor Final

En la parte del usuario/consumidor final (zona derecha de la Fig. 1), la plataforma incluye los módulos necesarios para implementar las siguientes funcionalidades:

- Sintonización, recepción y reproducción de canales DVB-T. Estas funcionalidades están integradas en el módulo MS.
- Descubrimiento de contenidos adicionales relacionados, accesibles vía *broadband*, accediendo e interpretando la información contenida el fichero XML mencionado con anterioridad. Esta funcionalidad está integrada en el MS.
- Descubrimiento de un MS activo por parte de un CS y la asociación entre ellos. Por un lado, se ha adoptado una implementación *open-source* del

protocolo DIAL (*Discovery And Launch Protocol*) [26], adoptado por HbbTV para descubrimiento automático de los dispositivos MS y CS involucrados. En particular, dicha implementación se ha integrado en el terminal híbrido, siendo activada al lanzar el MS y quedando a la espera de nuevas conexiones de CS. Por otro lado, las funcionalidades que ofrece el protocolo *Simple Service Discovery Protocol* (SSDP) [27] se han desarrollado desde cero. El proceso que corresponde al uso de SSDP es invocado cada vez que un CS quiere descubrir un dispositivo con un servidor DIAL activo. Tras el descubrimiento, ambos módulos/dispositivos (MS y CS) se asocian y se comunican a través de un canal bidireccional (vía *websockets*). Este canal de comunicación permite el intercambio de información relevante para conseguir el objetivo de la sincronización híbrida.

- Identificación, selección, recepción, procesado y reproducción (adaptativa) de los contenidos *broadband* relacionados. Estas funcionalidades están integradas en el módulo CS, que puede ser ejecutado en el terminal híbrido (mismo dispositivo que el MS) o en uno o varios dispositivos independientes, pero conectados a la misma red local (ver Fig. 1). La identificación y selección del contenido adicional *broadband* se logra gracias a la información disponible en el fichero XML, cuya dirección está insertada en el *TEMI location descriptor*, a su vez insertado en el flujo MPEG2-TS del contenido *broadcast*. El fichero XML es analizado por el MS, que extrae la información de interés necesaria para los CS y la envía a través del canal de comunicación establecido entre ambos. La recepción, procesado y reproducción de los contenidos adicionales *broadband* se consigue generando y enlazando los módulos de GStreamer necesarios para implementar un reproductor multimedia que soporte la naturaleza del contenido.
- Sincronización multimedia. La *sincronización intra-media* e *inter-media* para/entre los elementos multimedia (audio, video...) de cada flujo multiplexado recibido está soportada de manera nativa en GStreamer. La sincronización híbrida entre diferentes flujos en un mismo dispositivo (es decir, entre los contenidos reproducidos por el MS y el CS, ambos en el terminal híbrido), se consigue extrayendo las marcas temporales TEMI y comparando el estado de los procesos de reproducción de los flujos involucrados, utilizando para ello la información de las líneas de tiempo relativas (es decir, los valores de los campos PCR/PTS) y absolutas TEMI, y ajustando el estado de los procesos de reproducción de los flujos requeridos si la asincronía entre ellos excede un determinado umbral (configurable). Las líneas de tiempo se intercambian entre MS y CS a través del

canal bidireccional establecido previamente. Respecto a las técnicas de ajuste del proceso de reproducción, se pueden adoptar dos alternativas. Por una parte, se puede realizar técnicas de ajuste agresivas, como son saltos y pausas. Esta alternativa es apropiada cuando la asincronía supera un cierto umbral superior ($U1$) o cada vez que se active la reproducción de un nuevo contenido (o flujo) *broadband*. Por otro lado, la adopción de técnicas de reproducción multimedia adaptativa (*Adaptive Media Playout*, AMP) es mucho más conveniente cuando el nivel de asincronía que debe corregirse supera un umbral $U2$, pero no supera el umbral $U1$ ($U1 > U2$). AMP consiste en ajustar de manera suavizada (acelerando o reduciendo) la tasa de reproducción para corregir situaciones de asincronía. Permite alcanzar una mayor precisión en la sincronización y evitar discontinuidades en los procesos de reproducción, que pueden resultar molestas para la percepción humana (mala calidad de experiencia, QoE) [28]. Por su parte, IDES entre flujos reproducidos por el MS y por otros CS en diferentes dispositivos independientes se puede alcanzar de manera similar, aunque, en este caso, las funcionalidades de MS y CS se integran en diferentes dispositivos.

IV. EVALUACIÓN

El funcionamiento de la plataforma, para cada una de sus funcionalidades, ha sido evaluado, tanto objetiva como subjetivamente, obteniendo resultados muy satisfactorios. El escenario de evaluación, la metodología seguida y algunos de los resultados obtenidos se presentan a continuación.

A. Escenario evaluado

La plataforma se ha evaluado para el caso de uso de TV multi- vista y multi-dispositivo. Este caso de uso permite al proveedor de contenidos, para un determinado evento (ej. un evento deportivo, un concierto musical...), ofrecer un flujo *broadcast* (DVB) con las escenas seleccionadas por el realizador, complementadas con contenidos adicionales capturados por otras cámaras, ofreciendo un punto de vista diferente e información adicional del evento, que son distribuidos vía *broadband* (ej. DASH, HLS, RTP...). Los contenidos provenientes de estas cámaras adicionales pueden ser reproducidos por el CS, bien en el mismo terminal híbrido (ej. en modo PiP o mosaico, junto al contenido principal) o bien en dispositivos secundarios, de manera sincronizada. De acuerdo con un estudio realizado recientemente por nuestro grupo [29], se trata de un caso de uso que despierta un gran interés en los consumidores y que tiene mucho potencial comercial.

Concretamente, el escenario evaluado en este trabajo consiste en cuatro cámaras con distintas vistas

de un concierto¹. Las escenas seleccionadas por el realizador son distribuidas vía DVB-T y son reproducidas por el MS en el terminal híbrido. Por otro lado, los contenidos capturados por la misma cámara junto con los capturados por tres cámaras adicionales, se preparan y distribuyen vía DASH, pudiendo ser reproducidos por el CS en el mismo terminal híbrido o en dispositivos secundarios. La aplicación HbbTV desarrollada para el CS incluye un menú que permite la selección y/o cambio dinámico de la cámara a reproducir. Además, se ha configurado un umbral de ± 80 ms, de manera que no se realizan ajustes de reproducción si la asincronía calculada no lo supera.

La Fig. 2 muestra el aspecto del escenario multi- vista y multi-pantalla que se ha implementado y evaluado, el cual permite una mayor inmersión y personalización en las experiencias de consumo de contenidos de TV.

La Fig. 3, muestra un ejemplo del fichero XML utilizado para notificar la existencia de los contenidos híbridos relacionados en este escenario. Aunque en el escenario evaluado todos los contenidos *broadband* se distribuyen vía DASH, la figura muestra cómo se señalarían en el caso en el que se distribuyeran utilizando otras tecnologías *broadband*, como HLS y RTP.



Fig. 2 Escenario multi- vista y multi-dispositivo evaluado

```
<Media Content Description File>
<CAM id=1 protocol="dash" media_type="AV" media_format=
"h264/aac" metadata="TV Camera/english" uri=
"http://192.16.0.10/dash/cam1/cam1.mpd" />
<CAM id=2 protocol="hls" media_type="AV" media_format=
"h264/aac" metadata="Close-Up Camera/english" uri=
"http://192.16.0.10/hls/cam2/cam2.m3u8" />
<CAM id=3 protocol="rtp" media_type="AV" media_format=
"h264/aac" metadata="First Row Camera/english" uri=
"rtsp://224.0.0.10:5001/show" />
<URL id=1 protocol="http" media_type="website"
media_format="html" metadata="url_event/english" >
http://iim.webs.upv.es />
<LASTUPDATE format="dd/mm/yyyy-hh:mm:ss">10/11/2017-
15:45:30 />
<CLOCK protocol="ntp" media_type="time"
format="64_bit_ntp_time" url="ntp.upv.es" value=
390909493091340" />
</Media Content Description File >
```

Fig. 3: Ejemplo del fichero XML para el escenario evaluado

¹ Youtube es el propietario de los contenidos multimedia utilizados en el escenario de evaluación de este trabajo. Se trata de un concierto, grabado con cuatro cámaras (y con mismo audio), de la cantante Madilyn Bailey en Los Ángeles, 2015.

Vídeos demostrativos mostrando las capacidades y funcionalidades de la plataforma están disponibles en <http://iim.webs.upv.es/prototypes.html>

B. Evaluación objetiva

La plataforma permite la reproducción de contenidos *broadband* relacionados a través de los CS, tanto en el terminal híbrido como en dispositivos secundarios. Cuando se reproducen en el mismo terminal híbrido, los contenidos pueden presentarse en modo mosaico o PiP. Aunque los ecos del audio y la sincronización entre imagen y audio (*lip-sync*) puede percibirse, resulta difícil discernir visualmente el nivel de asincronía existente mediante estas configuraciones. Una manera sencilla de comprobar el nivel de sincronización consiste en reproducir el mismo contenido tanto en el MS como en el CS, pero recortando las imágenes recibidas por cada uno en su mitad derecha e izquierda y colocarlas una junto a la otra. Si la reproducción de ambos procesos se percibe como un único vídeo (ver Fig. 4 y vídeos disponibles en el enlace proporcionado) significa que la precisión conseguida es muy satisfactoria.

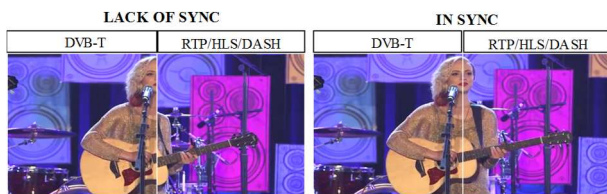


Fig. 4: Comprobación visual de la sincronización

Otro método utilizado para comprobar visualmente la asincronía entre los flujos multimedia consumidos consiste en superponer un texto indicando el número de trama del vídeo, insertado durante su proceso de codificación. Mediante este método, se puede tomar fotos (ver Fig. 2) o grabar vídeos del escenario evaluado (ver enlace proporcionado) e ir pausando el vídeo en instantes específicos para comprobar el desfase entre tramas existente en cada momento.

Además, se ha utilizado otro método más sistemático y transparente para los usuarios para medir las asincronías. Consiste en comparar (a nivel de código) las líneas temporales e instantes de reproducción de cada flujo e ir registrando las asincronías calculadas. Como ejemplo, la Fig. 5 muestra la evolución de la asincronía entre el MS y el CS (en dispositivos diferentes) cuando reproducen un flujo DVB-T y uno DASH, respectivamente. Puede observarse que, tras un proceso inicial de sincronización cuando se lanza el CS, la asincronía se mantiene dentro de límites aceptables (inferiores a ± 80 ms). Cuando se cambia a una cámara/vista diferente en el CS, se recupera la sincronización tras un periodo de ajuste inicial. La Fig. 6 confirma que el nivel de asincronía se mantiene dentro de los límites permitidos, exceptuando situaciones esporádicas (ej. al lanzar el CS, al cambiar de cámara, en situaciones de congestión de red o del dispositivo...). Por tanto, el

rendimiento de la plataforma en cuanto a precisión de sincronización es satisfactorio.

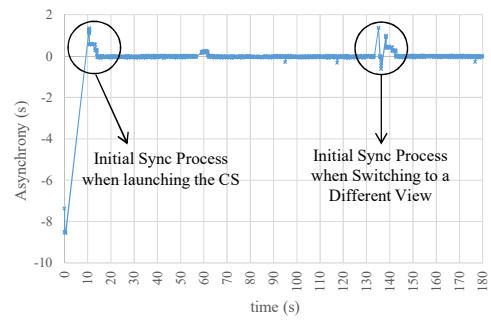


Fig. 5: Evolución de la asincronía entre el MS y el CS

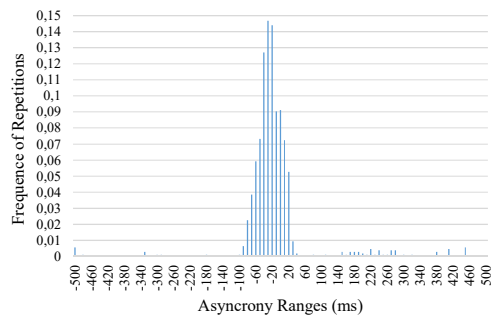


Fig. 6: Distribución de los valores de asincronía

C. Evaluación subjetiva

El rendimiento de la plataforma (en cuanto a sincronización, estabilidad y fluidez), para el caso de uso presentado, junto con la usabilidad de la misma, la percepción de los niveles de retardos y sincronización y el interés por parte de los consumidores en este tipo de plataformas y servicios, también han sido evaluados subjetivamente. Concretamente, 24 usuarios participaron en el estudio, los cuales, tras una breve introducción, probaron la plataforma utilizando dos tipos de dispositivos (*tablets* y *smartphones*).

En primer lugar, se presentó a los usuarios la visualización del mismo contenido tanto en el MS como en el CS, con el objetivo de determinar la precisión en la sincronización alcanzada. Para ello se mostraron 5 casos presentados de forma aleatoria, donde se forzaron asincronías fijas en el CS de 0, ± 1 y ± 3 segundos. Tras la visualización de cada caso, los usuarios utilizaron la métrica MOS (*Mean Opinion Score*) para evaluar el nivel de sincronización percibido, en la que un valor de 1 se corresponde con un nivel de sincronización muy malo y un 5 con un nivel excelente. Los resultados que se muestran en la Tabla 1 confirman que la QoE fue muy buena cuando no se forzó asincronía, aunque empeoró significativamente al aumentarla. Se presentan los valores obtenidos junto con el intervalo de confianza (IC) del 95% para cada caso visualizado.

Tabla 1: MOS \pm IC del 95% para cada caso

Caso	-3"	-1"	0"	+1"	+3"
MOS	1.75	2.17	4.63	2.46	1.75
\pm IC 95%	± 0.34	± 0.43	± 0.21	± 0.45	± 0.34

En segundo lugar, se permitió a los participantes utilizar libremente la plataforma, tras lo cual rellenaron un cuestionario. Los resultados obtenidos demuestran que, en general, los usuarios están muy interesados en este tipo de servicios multimedia híbridos (94%). Se mostraron muy satisfechos con la usabilidad de la plataforma (100%), con la utilidad de la misma (83%) y con su fluidez, estabilidad y nivel de sincronización conseguido (96%, aproximadamente, en todas ellas).

V. CONCLUSIONES Y TRABAJO FUTURO

Este artículo ha presentado una plataforma extremo-a-extremo compatible con la versión más reciente del estándar HbbTV (2.0.1) para la distribución y reproducción sincronizada de contenidos híbridos (*broadcast* y *broadband*) relacionados, tanto en un único dispositivo (ej. TV conectadas) como en múltiples dispositivos (ej. escenarios multi-pantalla). La plataforma ha sido evaluada para el caso de uso de TV multi-vista y multi-dispositivo, obteniendo resultados prometedores en cuanto a su rendimiento, usabilidad e interés despertados, reflejando así su potencial en el actual paradigma de consumo de contenidos multimedia.

Como trabajo futuro, se pretende minimizar los retardos, mejorar la precisión de sincronización y optimizar el rendimiento para mejorar la QoE. Además, con el fin de tener una plataforma basada completamente en GStreamer, se desarrollará un módulo insertor de descriptores TEMI con dicho *framework*. Asimismo, se extenderá la plataforma con el fin de incluir soporte para diferentes sistemas como Windows o iOS; además de para otros tipos de formatos y contenidos multimedia (ej. UHD, omnidireccionales, etc.), así como contenido en vivo. Finalmente, la plataforma se implementará y probará en sistemas reales de TV.

AGRADECIMIENTOS

El trabajo ha sido financiando por el MINECO y FEDER con ref. TEC2013-45492-R. Agradecer a Samsung Electronics Iberia S.A. la prestación de los dispositivos para el desarrollo y evaluación de la plataforma.

REFERENCIAS

- [1] A. Begen, et al, "Watching Video over the Web: Part 1: Streaming Protocols", IEEE Internet Computing, 15(2), pp. 54-63, April 2011.
- [2] H. Schulzrinne, et al, "RTP: A Transport Protocol for Real-Time Applications", IETF Standard, RFC 3550, July 2003.
- [3] Information Technology – Dynamic Adaptive Streaming over HTTP (DASH) – Part 1: Media presentation description and segment formats, ISO/IEC 23009-1, 2014. http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=65274
- [4] HbbTV 2.0.1 Specification, HbbTV Association Resource Library, <https://www.hbbtv.org/resource-library>, July 2016.
- [5] J. Le Feuvre, et al, "A Test Bed for Hybrid Broadcast Broadband Services". MediaSync Workshop 2015, Brussels (Belgium), June 2015.

- [6] R. van Brandenburg, et al, "Immersive second-screen experiences using hybrid media synchronization", MediaSync Workshop 2013, Nantes (France), October 2013.
- [7] F. Boronat et al, "Multimedia group and inter-stream synchronization techniques: A comparative study", Information Systems, 34(1), pp. 108–131, March 2009.
- [8] Z. Huang, et al, "Evolution of Temporal Multimedia Synchronization Principles: A Historical Viewpoint", ACM TOMCAPP, Vol. 9, No. 1s, Article 34, October 2013.
- [9] M. Oskar van Deventer, et al, "Standards for Multi-stream and Multi-device Media Synchronization", IEEE Communications Magazine, 54(3), pp. 16-21, March 2016.
- [10] L. Beloqui, et al, "Understanding Timelines Within MPEG Standards", IEEE Communications Surveys & Tutorials, 18(1), pp. 368-400, Firstquarter 2016.
- [11] L. Ehley, B. Furht, M. Ilyas, "Evaluation of multimedia synchronization techniques", International Conference on Multimedia Computing and Systems, pp. 514-519, May 1994.
- [12] M. Armstrong, et al, "Enabling and Enriching Broadcast Services by Combining IP and Broadcast Delivery", BBC Research White Paper WHP 185, September 2010.
- [13] S. Aoki, et al, "A New Transport Scheme for Hybrid Delivery of Content over Broadcast and Broadband". IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB) 2011, Fraunhofer (Germany), June 2011.
- [14] K. Evensen, et al, "Improving the Performance of Quality-Adaptive Video Streaming over Multiple Heterogeneous Access Networks", ACM MMSYS 2011, pp. 57-68, San Jose, California (USA), February 2011.
- [15] L. Beloqui, et al, "Interactive Multi-source Media Synchronisation for HbbTV". MediaSync Workshop 2012, Berlin (Germany), October 2012.
- [16] C. Concolato, et al, "Synchronized Delivery of Multimedia Content over Uncoordinated Broadcast Broadband Networks". ACM MMSYS 2012, Chapel Hill, North Carolina (USA), February 2012.
- [17] ISO/IEC 13818-1:2013/PDAM 6 Delivery of Timeline for External Data, 2013, <http://mpeg.chiariglione.org/standards/mpeg-2/systems/textisoiec-13818-12013pdam-6-delivery-timeline-external-data>
- [18] A. Veenhuizen, et al, "Frame accurate media synchronization of heterogeneous media sources in an HBB context", MediaSync Workshop 2012, Berlin (Germany), October 2012.
- [19] ETSI TS 102 823 v1.1.1 Digital Video Broadcasting (DVB); Specification for the carriage of synchronized auxiliary data in DVB transport streams, 2005. http://www.etsi.org/deliver/etsi_ts/102800_102899/102823/01.01.01_60/ts_102823v010101p.pdf
- [20] GStreamer: open source multimedia framework. <https://GStreamer.freedesktop.org/>
- [21] J. Le Feuvre, C. Concolato, "Hybrid Broadcast Services using MPEG DASH", MediaSync Workshop 2013, Nantes (France), October 2013.
- [22] GPAC <https://gpac.wp.mines-telecom.fr/>
- [23] C. Howson et al, "Second Screen TV Synchronization", IEEE International Conference on Consumer Electronics, Berlin (Germany), January 2011.
- [24] R. van Brandenburg, et al, "Immersive second-screen experiences using hybrid media synchronization", MediaSync Workshop 2013, Nantes (France), October 2013.
- [25] D. Gómez, et al, "End-to-End DASH Platform including a Network-based and Client-based Adaptive Quality Switching Module", ACM MMSYS 2016, Klagenfurt (Austria), May 2016.
- [26] DIAL (Discovery And Launch) protocol specification, Version 1.7.2, 2015. <http://www.dial-multiscreen.org/dial-protocol-specification/DIAL-2ndScreenProtocol-1.7.2.pdf>
- [27] Simple Service Discovery Protocol SSDP, Internet Draft <http://quimby.gnus.org/internet-drafts/draft-cai-ssdp-v1-03.txt>
- [28] M. Montagud et al, "How to perform AMP? Cubic adjustments for improving the QoE", Computer Communications, Volume 103, pp 61-63, ISSN 0140-3664, May 2017.
- [29] F. Boronat, et al, "Preferencias, necesidades y expectativas de los usuarios españoles en escenarios multimedia híbridos broadcast/broadband", VI Congreso de TV Digital Interactiva (TVDi), Mallorca (Spain), October 2015.