



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Especificación, implementación y evaluación de soluciones de sincronización multimedia híbrida (broadcast/broadband) e inter-destinatario para posibilitar nuevas experiencias televisivas enriquecidas, personalizadas, inmersivas y compartidas.

Departamento de Comunicaciones

Universitat Politècnica de València

Tesis presentada para la obtención del grado de

Doctor por la Universitat Politècnica de València

Gandia, julio 2022

AUTOR:

Daniel Marfil Reguero

DIRECTOR:

Dr. Fernando Boronat Seguí

RESUMEN

Actualmente, existe una gran variedad de tecnologías de distribución de contenidos multimedia, de dispositivos de consumo, así como de tipos de contenidos. En cuanto a tecnologías de distribución, dichos contenidos se pueden enviar bien a través de redes broadcast, como las redes DVB (*Digital Video Broadcasting*), o bien a través de redes broadband, como las redes IP (por ejemplo, a través de Internet).

La coordinación y convergencia entre los dos tipos de tecnologías para distribuir contenidos relacionados ofrece un gran potencial a los proveedores de contenidos, así como grandes beneficios a los consumidores (una prueba de ello es el estándar Hybrid Broadcast Broadband TV o HbbTV). Sin embargo, la recepción de contenidos relacionados procedentes de diferentes redes híbridas (broadcast/broadband) conlleva una variabilidad de retardos significativa, por lo que se deben superar una serie de retos importantes, como, por ejemplo, la necesidad de mecanismos que garanticen un consumo sincronizado de dichos contenidos, bien en un mismo dispositivo o bien en distintos dispositivos (locales o remotos).

Además, por un lado, el paradigma de consumo de contenidos de TV está evolucionando, a nivel global, desde un consumo aislado y pasivo hacia nuevas experiencias interactivas y compartidas entre grupos de usuarios (incluso geográficamente dispersos), enriqueciendo así la experiencia de consumo de TV, facilitando la interacción y el contacto social entre los mismos. Un ejemplo típico es el caso de uso de TV Social, en el cual un grupo de consumidores, geográficamente distribuidos, están viendo la TV de forma conjunta, mientras interactúan socialmente vía herramientas de chat o audio/videoconferencia (caso de uso también conocido como "*watching apart together*"). Este nuevo paradigma es fruto de la necesidad de los humanos, como seres sociales que son, de compartir información y consumir contenido, formando así parte de un grupo.

Por otro lado, la comunidad científica y la industria están trabajando en el desarrollo de sistemas multimedia multisensoriales capaces de integrar efectos sensoriales, como aromas, efectos ambientales (iluminación especial, temperatura, humedad, humo, vaporizaciones...), etc., cuyo objetivo es estimular otros sentidos, más allá de la vista y el oído, y, por tanto, proporcionar experiencias multimedia más inmersivas y realistas.

Esta Tesis Doctoral centra su investigación en los diferentes tipos de mecanismos de sincronización multidispositivo y multiusuario, con el objetivo de conseguir un entorno inmersivo de consumo de contenidos de TV híbridos, *multimedia* (*multisensorial media*) y multiusuario.

Se han especificado, implementado y evaluado (tanto objetiva como subjetivamente) soluciones de sincronización de los procesos de reproducción de contenidos híbridos (prevenientes de redes broadcast y/o broadband) tanto en entornos multidispositivo, entre dispositivos locales (*Inter-Device Synchronization* o IDES), como en entornos multiusuario, entre destinatarios distribuidos en red (*Inter-Destination Media Synchronization* o IDMS), incluyendo la sincronización de efectos multisensoriales.

Para poder evaluarlas, se ha diseñado e implementado una plataforma extremo-a-extremo para la generación, distribución (adaptativa) y consumo de contenidos híbridos relacionados, tanto en un mismo dispositivo como en múltiples dispositivos locales (por ejemplo, escenario

multipantalla) o remotos (por ejemplo, TV Social). Se han generado y preparado contenidos oportunos (audiovisuales y multisensoriales), para ser distribuidos mediante dicha plataforma, con el fin de evaluar su funcionamiento, así como el de las soluciones de sincronización desarrolladas e implementadas en la misma. Dichos contenidos consisten en contenidos multimedia tradicionales (2D) y omnidireccionales (360), enriquecidos con efectos multisensoriales que mejoran el grado de realismo y la inmersividad de los usuarios. La plataforma es compatible con el estándar HbbTV y, además, incluye soluciones eficientes para aspectos clave que no están especificados en dicho estándar, pero que son necesarios para la implementación satisfactoria de servicios de TV híbridos y multidispositivo en la misma, como, por ejemplo: la necesidad de un buffer para el contenido broadcast (existente y opcional hasta la especificación HbbTV 2.0.2); mecanismos de señalización, asociación y descripción de contenido disponible y relacionado con el contenido principal; mecanismos de coordinación e interacción entre los dispositivos de consumo; o soluciones de sincronización para alinear en el tiempo el consumo de contenidos de forma satisfactoria.

Por un lado, los resultados obtenidos en la evaluación realizada muestran que el nivel de sincronización adquirido entre los procesos de reproducción de los diversos tipos de contenido disponibles a través de distintas tecnologías de transmisión alcanza unos valores suficientes para proporcionar al usuario un consumo con un nivel de asincronías imperceptible. Por otro lado, dicha precisión alcanzada ha resultado en una QoE para el consumidor muy satisfactoria, ofreciéndole al mismo un completo ecosistema de consumo de contenidos que incluye tanto contenido multimedia tradicional, como contenido omnidireccional y mulsemédia.

RESUM

Actualment, existeix una gran varietat de tecnologies de distribució de continguts multimèdia, de dispositius de consum, així com de diferents tipus de continguts. Respecte a tecnologies de distribució, aquests continguts es poden enviar bé a través de xarxes broadcast, com ara les xarxes DVB (Digital Video Broadcasting), o bé a través de xarxes broadband, com les xarxes IP (per exemple, a través d'Internet).

La coordinació i la convergència entre els dos tipus de tecnologies per distribuir continguts relacionats, ofereix un gran potencial als proveïdors de continguts, així com grans beneficis als consumidors (una prova d'això és l'estàndard Hybrid Broadcast Broadband TV o HbbTV). Tot i això, la recepció de continguts relacionats procedents de diferents xarxes híbrides (broadcast/broadband) comporta una variabilitat de retards significativa, per la qual cosa s'han de superar una sèrie de reptes importants, com és la necessitat de mecanismes que garanteixin un consum sincronitzat d'aquests continguts, bé en un mateix dispositiu o bé en diferents dispositius (locals o remots).

A més, per una banda, el paradigma de consum multimèdia evoluciona, a nivell global, des d'un consum aïllat i passiu de continguts de TV cap a noves experiències interactives i compartides entre grups d'usuaris (fins i tot geogràficament dispersos), enriquint així l'experiència de consum de TV, facilitant la interacció i el contacte social entre aquests. Un exemple típic és el concepte de TV Social, en el qual un grup de consumidors, geogràficament distribuïts, estan veient la TV de forma conjunta, mentre interactuen socialment via eines de xat o àudio/videoconferència (cas d'ús també conegut com a "*watching apart together*"). Aquest nou paradigma és fruit de la

necessitat dels humans, com a éssers socials que són, de compartir informació i consumir contingut, formant part d'un grup.

D'altra banda, la comunitat científica i la indústria estan treballant en el desenvolupament de sistemes multimèdia multisensorials capaços de reproduir efectes sensorials, com ara aromes, efectes ambientals (il·luminació especial, temperatura, humitat, fum, vaporitzacions...), etc., amb l'objectiu d'estimular altres sentits, més enllà de la vista i l'oïda, i, per tant, proporcionar experiències multimèdia més immersives i realistes.

Aquesta Tesi Doctoral centra la seua recerca en els diferents tipus de mecanismes de sincronització multidispositiu i multiusuari, amb l'objectiu d'aconseguir un entorn immersiu de consum de continguts híbrids, mulsemèdia (*multisensorial media*) i multiusuari.

S'han especificat, implementat i avaluat (tant objectivament com subjectivament) solucions de sincronització dels processos de reproducció de continguts híbrids (provinents de xarxes broadcast i/o broadband) tant en entorns multidispositiu, entre dispositius locals o propers (Inter-Device Synchronization o IDES), com en entorns multi-usuari, entre destinataris distribuïts en xarxa (Inter-Destination Media Synchronization o IDMS), incloent-hi la sincronització d'efectes multisensorials.

Per poder avaluar aquestes solucions de sincronització, s'ha dissenyat i implementat una plataforma extrem a extrem, per a la generació, distribució (adaptativa) i consum de continguts híbrids relacionats, tant en un mateix dispositiu com en múltiples dispositius locals (p. ex., un escenari multipantalla) o remots (p. ex., TV Social). S'han generat i preparat continguts oportuns (audiovisuals i multisensorials), per ser distribuïts mitjançant la plataforma desenvolupada, per tal d'avaluar-ne el funcionament, així com el de les solucions de sincronització desenvolupades i implementades en aquesta. Aquests continguts són continguts multimèdia tradicionals (2D) i omnidireccionals (360), enriquits amb efectes multisensorials que milloren el grau de realisme i la immersivitat dels usuaris. La plataforma és compatible amb l'estàndard HbbTV i, a més, inclou solucions eficients per a aspectes clau que no estan especificats en aquest estàndard, però que són necessaris per a la implementació satisfactòria de serveis de TV híbrids i multidispositiu en aquesta, com per exemple: la necessitat d'un buffer per al contingut broadcast (existent i opcional fins a l'especificació HbbTV 2.0.2); mecanismes de senyalització, associació i descripció de contingut disponible i relacionat amb el contingut principal; mecanismes de coordinació i interacció entre els dispositius de consum; o solucions de sincronització per alinear en el temps el consum de continguts de manera satisfactòria.

D'una banda, els resultats obtinguts a l'avaluació realitzada mostren, per una part, que el nivell de sincronització adquirit entre els processos de reproducció dels diversos tipus de contingut disponibles a través de diferents tecnologies de transmissió assoleix uns valors suficients per proporcionar a l'usuari un consum amb un nivell d'asincronies imperceptible. D'altra banda, aquesta precisió aconseguida ha resultat en una QoE per a l'usuari molt satisfactòria, oferint al consumidor un complet ecosistema de consum de continguts que inclou tant contingut multimèdia tradicional com contingut omnidireccional i mulsemèdia.

ABSTRACT

Currently, there is a wide variety of multimedia content distribution technologies, consumer devices, as well as types of content. In terms of distribution technologies, such content can be delivered either over broadcast networks, such as DVB (Digital Video Broadcasting) networks, or over broadband networks, such as IP networks (e.g., over the Internet).

The coordination and convergence between the two types of technologies to distribute related content offers great potential to content providers, as well as great benefits to consumers (a proof of this is the Hybrid Broadcast Broadband TV or HbbTV standard). However, the reception of related content from different hybrid networks (broadcast/broadband) entails significant delay variability, so a number of important challenges must be overcome, such as the need for mechanisms to ensure synchronized consumption of such content, either on the same device or on different devices (local or remote).

Moreover, on the one hand, the paradigm of multimedia consumption is evolving, globally, from an isolated and passive consumption of TV content to new interactive and shared experiences among groups of users (even geographically dispersed), thus enriching the TV consumption experience, facilitating interaction and social contact among them. A typical example is the concept of Social TV, in which a group of consumers, geographically distributed, are watching TV together, while interacting socially via chat or audio/video conferencing tools (use case also known as "watching apart together"). This new paradigm is the result of the need of humans, as social beings, to share information and consume content, thus being part of a group.

On the other hand, the scientific community and industry are working on the development of multi-sensory multimedia systems capable of reproducing sensory effects, such as scents, environmental effects (special lighting, temperature, humidity, smoke, vaporizations...), etc., whose objective is to stimulate other senses, beyond sight and hearing, and, therefore, provide more immersive and realistic multimedia experiences.

This PhD focuses its research on the different types of multidevice and multiuser synchronization mechanisms, with the aim of achieving an immersive environment of hybrid, multimedia (multisensory media) and multiuser consumption.

Synchronization solutions for hybrid content reproduction processes (broadcast and/or broadband networks) have been specified, implemented, and evaluated (both objectively and subjectively) in multidevice environments, between local or nearby devices (Inter-Device Synchronization or IDES), and in multiuser environments, between distributed network receivers (Inter-Destination Media Synchronization or IDMS), including the synchronization of multisensory effects.

In order to evaluate them, an end-to-end platform has been designed and implemented for the generation, (adaptive) distribution and consumption of related hybrid contents, both on a single device and on multiple devices locally (i.e., multiscreen scenario) or remotely (e.g., Social TV). Appropriate contents (audiovisual and multisensory) have been generated and prepared to be distributed through the developed platform, in order to evaluate its performance, as well as that of the synchronization solutions developed and implemented in it. These contents are traditional (2D) and omnidirectional (360) multimedia contents, enriched with multisensory effects that

improve the degree of realism and immersiveness of the users. The platform is compatible with the HbbTV standard and, in addition, includes efficient solutions for key aspects that are not specified in that standard, but are necessary for the successful implementation of hybrid and multi-device TV services on it, such as: the need for a buffer for broadcast content (existing and optional until HbbTV 2.0.2); mechanisms for signaling, association and description of content available and related to the main content; coordination and interaction mechanisms between consumption devices; or synchronization solutions to align in time the consumption of content in a satisfactory way.

On the one hand, the results of the evaluation carried out show, firstly, that the level of synchronization acquired between the playout processes of the different types of content available through different transmission technologies reaches sufficient values to provide the user with a consumption with an imperceptible level of asynchronies. On the other hand, that the precision achieved results in a QoE for the user has been very satisfactory for the user. Thus, providing the users with a complete environment for multimedia and mulsemmedia content consumption.

AGRADECIMIENTOS

A Fernando Boronat, por todos sus consejos y todo el soporte que me ha dado a lo largo de este camino. Al grupo de investigación Immersive Interactive Media por toda la ayuda que me han dado durante todo este tiempo.

A mi familia, en especial a mi madre por su paciencia y su confianza ciega en mi y a Sara, por su amor infinito.

TABLA DE CONTENIDO

Índice de figuras	17
Índice de tablas	23
Índice de ecuaciones	25
Acrónimos	27
Capítulo 1. Introducción y objetivos de la Tesis Doctoral	29
1.1 Introducción y marco de la Tesis Doctoral.....	29
1.2 Objetivos de la Tesis Doctoral.....	34
1.3 Metodología de la Tesis Doctoral.....	35
1.4 Organización de la memoria.....	36
Capítulo 2. Conceptos y mecanismos relacionados	39
2.1 Tipos de sincronización multimedia	39
2.2 Entrega híbrida de contenidos	39
2.3 Problema de la variabilidad de retardos en escenarios híbridos.....	40
2.4 Necesidad de sincronización híbrida (<i>Hybrid Synchronization</i>).....	41
2.5 Sincronización de relojes	42
2.6 Inserción de líneas temporales en el flujo broadcast	43
2.7 Arranque y adquisición de contenidos híbridos	46
2.8 Descubrimiento y comunicación entre dispositivos en un entorno local	47
2.8.1 Discovery and Launch Protocol	48
2.8.2 Companion Screens and Streams.....	48
2.9 Estándares que contemplan la recepción y el consumo de contenidos híbridos	49
2.9.1 Hybrid Broadcast Broadband TV	49
2.9.2 Multimedia Home Platform.....	50
2.9.3 MPEG Media Transport	50
2.9.4 Ginga	50
2.9.5 Advanced Television Systems Committee	51
2.10 Componentes de las soluciones de sincronización entre flujos multimedia	51
2.10.1 Esquemas de control de la sincronización.....	51
2.10.2 Estrategias de selección de la referencia de sincronización.....	52
2.10.3 Técnicas de ajuste en el proceso de reproducción del contenido	53
2.11 Contenidos de nueva generación	53
2.11.1 Contenido omnidireccional.....	53
2.11.2 Contenido multisensorial	56

2.12	Conclusiones	57
Capítulo 3.	Trabajos relacionados	59
3.1	Sincronización híbrida.....	59
3.1.1	Sincronización de la reproducción de contenidos híbridos en un único dispositivo (terminal híbrido).....	59
3.1.2	Sincronización de la reproducción de contenidos híbridos en varios dispositivos locales 61	
3.1.3	Sincronización de la reproducción de contenidos híbridos en varios dispositivos remotos 62	
3.2	TV Social.....	63
3.3	Vídeo omnidireccional	65
3.4	Mulsemmedia	66
3.4.1	Frameworks, prototipos y plataformas mulsemmedia.....	68
3.4.2	Transmisión de contenidos mulsemmedia a través de redes broadcast	70
3.4.3	Sincronización de contenido mulsemmedia	71
3.5	Conclusiones	71
Capítulo 4.	Soluciones de sincronización multimedia y multisensorial híbrida IDES e IDMS	75
4.1	Introducción	75
4.2	Tecnologías, estándares y protocolos adoptados.....	76
4.2.1	Transmisión híbrida de contenidos	76
4.2.2	Señalización de efectos multisensoriales.....	76
4.2.3	Descubrimiento y comunicación entre dispositivos.....	77
4.3	Preparación del contenido	78
4.3.1	Inserción de marcas temporales en el contenido principal.....	78
4.3.2	Señalización de la existencia de contenidos complementarios vía redes broadband 78	
4.4	Técnicas de ajuste del proceso de reproducción	80
4.5	Diseño y especificación de una solución de sincronización de la reproducción de contenidos híbridos en uno o varios dispositivos locales	81
4.6	Diseño y especificación de una solución de sincronización de la reproducción de contenidos híbridos en varios dispositivos remotos	85
4.6.1	Mensajes de sincronización inter-destinatario para el control de la sesión y de la sincronización.....	87
4.6.2	Creación de una sesión de sincronización inter-destinatario	90
4.6.3	Unirse a una sesión de sincronización inter-destinatario	91
4.6.4	Abandonar una sesión de sincronización inter-destinatario.....	91

4.6.5	Solución de sincronización de sincronización inter-destinatario	92
4.7	Diseño y especificación de una solución de sincronización de la reproducción de contenidos híbridos y la generación de efectos multisensoriales	98
4.7.1	Controlador mulsemmedia.....	99
4.7.2	Uso de una interfaz de programación de aplicaciones REST	100
4.7.3	Gestión y sincronización del contenido mulsemmedia.....	101
4.8	Conclusiones	102
Capítulo 5. Plataforma extremo-a-extremo para la generación, distribución y consumo sincronizado de contenidos híbridos, incluyendo contenidos omnidireccionales y efectos multisensoriales		103
5.1	Introducción	103
5.2	Arquitectura de la plataforma	103
5.2.1	Módulo de sincronización de relojes y referencia temporal global	107
5.2.2	Módulos de sincronización multimedia.....	107
5.3	Parte del proveedor de contenidos	110
5.3.1	Módulos de generación de contenidos.....	110
5.3.2	Módulo de generación del fichero Related Media Content File	114
5.3.3	Módulos de transmisión de contenidos	114
5.4	Parte del consumidor o usuario final.....	115
5.4.1	Terminal híbrido – Módulo Main Screen	115
5.4.2	Módulo Companion Screen	118
5.5	Conclusiones	130
Capítulo 6. Evaluación objetiva y resultados		131
6.1	Introducción	131
6.2	Escenario inter-dispositivo básico.....	131
6.2.1	Escenario de evaluación	132
6.2.2	Metodologías de evaluación empleadas	136
6.2.3	Resultados obtenidos.....	137
6.2.4	Discusión.....	140
6.3	Escenario inter-dispositivo inmersivo	141
6.3.1	Escenario de evaluación	142
6.3.2	Metodología de evaluación empleada	144
6.3.3	Resultados obtenidos.....	144
6.3.4	Discusión.....	147
6.4	Escenario inter-destinatario.....	148
6.4.1	Escenario de evaluación	148

6.4.2	Metodología de evaluación empleada	151
6.4.3	Resultados obtenidos.....	151
6.4.4	Discusión.....	161
6.5	Conclusiones	162
Capítulo 7.	Evaluación subjetiva y resultados	163
7.1	Introducción	163
7.2	Directrices comunes en todas las evaluaciones realizadas por los participantes 163	
7.3	Escenario inter-dispositivo básico.....	164
7.3.1	Participantes	165
7.3.2	Metodología de evaluación.....	165
7.3.3	Resultados obtenidos.....	166
7.3.4	Discusión.....	171
7.4	Escenario inter-dispositivo inmersivo	171
7.4.1	Participantes	172
7.4.2	Metodología de evaluación.....	173
7.4.3	Resultados obtenidos.....	173
7.4.4	Discusión.....	180
7.5	Escenario inter-destinatario.....	180
7.5.1	Interfaz de mensajes contextualizados por evento en el canal de chat	182
7.5.2	Participantes	183
7.5.3	Metodología de evaluación.....	183
7.5.4	Resultados obtenidos.....	184
7.5.5	Discusión.....	192
7.6	Escenario mulsemidia.....	192
7.6.1	Participantes	193
7.6.2	Metodología de evaluación.....	194
7.6.3	Resultados obtenidos.....	195
7.6.4	Discusión.....	199
7.7	Conclusiones	199
Capítulo 8.	Conclusiones y trabajo futuro	201
8.1	Conclusiones	201
8.2	Trabajo futuro	204
Referencias		207
Anexo	Publicaciones y trabajos derivados de la Tesis Doctoral	219

Publicaciones en revistas de investigación	219
Publicaciones en congresos/workshops de investigación.....	220
Capítulos de libro.....	220

ÍNDICE DE FIGURAS

Fig. 1 Escenario de entrega de contenidos multimedia híbridos	40
Fig. 2 Variabilidad de retardos extremo-a-extremo	40
Fig. 3 Mecanismo de inserción de marcas temporales especificado por la especificación ETSI TS 102 823.....	44
Fig. 4 Dos opciones de inserción de los descriptores Timing and External Media Information ...	45
Fig. 5 Inserción de los descriptores TEMI en el campo de adaptación de los paquetes de transporte	46
Fig. 6 Necesidad de descubrimiento y comunicación en un escenario mono y multidispositivo.	47
Fig. 7 Esquemas de control del proceso de sincronización entre procesos de reproducción	52
Fig. 8 Porción visualizada por el usuario en un contenido omnidireccional	54
Fig. 9 Viraje (yaw), inclinación (pitch) y balanceo (roll), además de la traslación en los ejes X, Y, Z.	55
Fig. 10 Proyecciones más comunes en vídeo 360: ERP (izq.), CMP (centro) y piramidal (dch.).....	55
Fig. 11 Ejemplo de escenario de consumo de contenido mulsemedia (adaptación de [52])	56
Fig. 12 SEVino 2.0 tool [121]	69
Fig. 13 Interfaz del simulador SESim	69
Fig. 14 Tecnologías de transmisión adoptadas en esta Tesis Doctoral	76
Fig. 15 ejemplo de fichero SEM	77
Fig. 16 Componentes y funcionalidades de comunicación involucrados en un entorno local multidispositivo basado en HbbTV y DIAL.....	78
Fig. 17 Ejemplo de fichero RMCF	80
Fig. 18 Diagrama de flujo del dispositivo principal ejecutando la MS app	82
Fig. 19 Campos del mensaje de sincronización IDES.....	83
Fig. 20 Relación entre las marcas temporales utilizadas en el mecanismo de sincronización propuesto para escenarios mono- y multidispositivo (locales)	83
Fig. 21 Diagrama de flujo de dispositivo ejecutando la Companion Screen app.....	84
Fig. 22 Esquemas de control de sincronización en el escenario inter-destinatario (centralizado en la izquierda y distribuido en la derecha)	86
Fig. 23 Módulos integrados dentro del IDMS Manager	86
Fig. 24 Formato genérico de los mensajes IDMS propuestos.....	87
Fig. 25 Formato del mensaje IDMS Create Session ACK.....	88
Fig. 26 Formato del mensaje IDMS Join	88
Fig. 27 Formato de los ítems incluidos dentro del mensaje IDMS Join.....	88
Fig. 28 Formato del mensaje IDMS Leave	89
Fig. 29 Formato del mensaje IDMS Report.....	89
Fig. 30 Formato del mensaje IDMS Settings.....	89
Fig. 31 Diagrama de flujo seguido por un cliente IDMS para crear una sesión IDMS bajo un esquema de control DCS	91
Fig. 32 Diagrama de flujo del proceso seguido en el IDMS Manager en el esquema de control SMS	93
Fig. 33 Diagrama de flujo de un Cliente IDMS bajo un esquema de control SMS	94
Fig. 34 Diagrama de flujo de un Cliente IDMS bajo un esquema de control DCS.....	95

Fig. 35 Situación normal (superior) y no-deseada (inferior)	96
Fig. 36 Fichero SEM con el atributo global-timestamp.....	98
Fig. 37 Proceso de asociación de un valor basado en NTP a partir de un valor PTS	99
Fig. 38 Descubrimiento del MC a través de DIAL [30]	99
Fig. 39 Ejemplo de una petición a un recurso disponible (izquierda) o no disponible (derecha) entre la MS/CS app y el MC.....	101
Fig. 40 Diagrama de flujo de las principales tareas para la activación de contenidos mulsemmedia por parte de las apps MS o CS.....	102
Fig. 41 Visión general/arquitectura de la plataforma desarrollada.	104
Fig. 42 Funcionalidades incluidas en la parte del proveedor de contenidos.....	104
Fig. 43 Parte del consumidor final en un escenario IDMS.	105
Fig. 44 Funcionalidades involucradas en la parte del consumidor final (MS app), incluyendo los módulos necesarios para un escenario IDMS.	106
Fig. 45 Funcionalidades involucradas en la parte del consumidor final (MS app), incluyendo los módulos necesarios para un escenario IDMS.....	106
Fig. 46 Funcionalidades involucradas en la parte del consumidor final (CS app), incluyendo los módulos necesarios para un escenario IDMS.	107
Fig. 47 Relación entre el IDMS Manager y el servidor de chat en un escenario IDMS	109
Fig. 48 Herramienta SeVino2 para la inserción de efectos multisensoriales	113
Fig. 50 Software StreamXpress y tarjeta moduladora DTA-2111, Dektec	114
Fig. 51 Componentes GStreamer para la integración de un reproductor de contenido DVB-T.	116
Fig. 52 Código implementado para la configuración del receptor DVB-T.	117
Fig. 53 Código implementado como callback para recibir la información insertada en el descriptor TEMI	117
Fig. 54 Código desarrollado para lanzar una petición y recibir una respuesta, ambas multicast	120
Fig. 55 Componentes GStreamer para la integración de un reproductor de contenido adaptativo	120
Fig. 56 Ubicación del contenido audiovisual (gris claro) en el salón virtual (izquierda) y en el reproductor omnidireccional (derecha).....	122
Fig. 57 Aspecto del escenario inmersivo del salón virtual visualizado en una pantalla tradicional (a) y en un HMD (b)	122
Fig. 58 Aspecto final del reproductor omnidireccional en una pantalla tradicional (a) y en un HMD (b).....	123
Fig. 59 Selección del GUI del reproductor omnidireccional en función del dispositivo que lo ejecuta	123
Fig. 60 Tecnologías utilizadas en el desarrollo del escenario inmersivo	124
Fig. 61 Microcontrolador NodeMCU.....	125
Fig. 62 Diagrama de flujo que muestra el comportamiento del MC.....	126
Fig. 63 Esquema (izquierda) y vista parcial del circuito electrónico (derecha) encargado de actuar sobre los dispositivos encargados de generar los efectos sensoriales de viento o calor.	126
Fig. 64 Controlador PWM de motores DC.....	127
Fig. 65 Esquema (izquierda) y vista parcial del circuito electrónico (derecha) encargado de activar los ventiladores del dispositivo generador aromas	127
Fig. 66 Aspecto del filtro paso bajo.....	128
Fig. 67 Croquis del circuito amplificador de la señal	128

Fig. 68 Esquema (izquierda) y vista parcial del circuito electrónico (derecha) encargado de la generación de vibraciones a partir de un subwoofer de baja frecuencia	128
Fig. 69 Ejemplo de división del borde del contenido en regiones para obtener el valor de color promedio.....	129
Fig. 70 Proceso de generación (arriba) y vista parcial del circuito implementado para la emisión de la luz ambiental coherente con el contenido presentado en el Main Screen	130
Fig. 71 Escenario inter-dispositivo implementado en la plataforma para la evaluación objetiva de la solución IDES básica (sin incluir HMD)	132
Fig. 72 Aspecto de las aplicaciones MS y CS desarrolladas para la plataforma, tanto en el terminal híbrido como en el dispositivo secundario	133
Fig. 73 Menú emergente para seleccionar el contenido broadband disponible en MS app (superior) y CS app (inferior)	133
Fig. 74 Escenario IDES básico formado por undispositivo principal (ejecutando la MS app) y 4 dispositivos secundarios (ajactuando la CS app).....	134
Fig. 75 Fichero RMCF utilizado para el Escenario básico IDES	135
Fig. 76 Mismo contenido broadcast y broadband cortado por la mitad y reproducido por aplicaciones MS (parte izquierda del contenido) y CS (parte derecha del mismo) para comprobar visualmente el nivel de sincronización alcanzado en el terminal híbrido	136
Fig. 77 Valores de asincronía registrados por la CS app durante una sesión de la evaluación objetiva y la distribución de los valores registrados.....	140
Fig. 78 Escenario ides implementado en la plataforma para la evaluación objetiva de la solución ides inmersiva (incluyendo HMD)	142
Fig. 79 Aspecto del salón virtual visto en pantalla plana (izquierda superior) y en HMD (izquierda inferior). Aspecto del reproductor omnidireccional visto en pantalla plana (derecha superior) y en HMD (derecha inferior).....	143
Fig. 80 Etiquetas y parámetros específicos para señalar la disponibilidad del contenido omnidireccional en el fichero RMCF	143
Fig. 81 Valores de asincronía registrados por cada dispositivo secundario y número de coincidencias durante la sesión	145
Fig. 82 Función de frecuencia acumulada de los valores absolutos de las asincronías registradas en cada dispositivo secundario.....	146
Fig. 83 Valores de asincronía registrados por los dispositivos secundarios durante una sesión de 10 minutos.....	147
Fig. 84 Valores de asincronía y ajustes registrados en el dispositivo secundario smartphone Samsung Galaxy S5 durante una sesión adicional de 10 minutos (se incluye una media móvil tomando grupos de 10 muestras).....	147
Fig. 85 Visión general del escenario inter-destinatario.....	148
Fig. 86 Ejemplo de escenario IDMS con dos destinatarios independientes.	149
Fig. 87 Fichero RMCF utilizado en el escenario IDMS.....	151
Fig. 88 Número medio de ajustes necesarios durante la evaluación IDMS - SMS con 4 destinatarios involucrados	152
Fig. 89 Valor medio del RMS e I.C. 95% de los valores de la asincronía registrados en las 10 sesiones realizadas, obtenidos en cada MS app y en cada caso, respecto a la solución de sincronización IDMS con un esquema de control DCS.....	154

Fig. 90 Número medio de ajustes realizados durante la evaluación IDMS - DCS con 4 destinatarios involucrados	154
Fig. 91 Número medio de ajustes realizados durante la evaluación IDMS - DCS con 4 destinatarios involucrados	154
Fig. 92 Valor medio del RMS e I.C. 95% de los valores de la asincronía registrados en las 10 sesiones realizadas, obtenidos en cada CS app durante la evaluación IDMS.....	156
Fig. 93 Función de frecuencia acumulada de los valores absolutos de las asincronías registradas durante la evaluación objetiva en cada dispositivo secundario para cada uno de los escenarios IDMS evaluados (con esquemas de control SMS y DCS).....	157
Fig. 94 Comportamiento de los dispositivos secundarios en un escenario IDMS – SMS con referencia de sincronización el punto de reproducción de la MS app del dispositivo principal más atrasado.....	158
Fig. 95 Comportamiento de los CS en un escenario IDMS – SMS con referencia de sincronización el punto promedio de la reproducción de las apps MS de los 4 dispositivos principales involucrados	159
Fig. 96 Comportamiento de los dispositivos secundarios en un escenario IDMS – DCS con referencia de sincronización el punto de reproducción de la MS app del dispositivo principal más atrasado.....	160
Fig. 97 Comportamiento de los dispositivos secundarios en un escenario IDMS – DCS con referencia de sincronización el punto promedio de la reproducción de las apps MS de los 4 dispositivos principales involucrados.....	161
Fig. 98 Valor MOS de los cinco casos de asincronía forzada indicados en la Tabla 33	167
Fig. 99 Uso de dispositivos secundarios mientras se consume contenido de TV.....	167
Fig. 100 Percepción del nivel de sincronización alcanzado entre los procesos de reproducción de los dispositivos principal y secundario	168
Fig. 101 Aceptabilidad del retardo inicial al lanzar la CS app en el dispositivo secundario.....	169
Fig. 102 Aceptabilidad del retardo al cambiar de cámara en la CS app en el dispositivo secundario	169
Fig. 103 Interés y utilidad del escenario evaluado.....	169
Fig. 104 Facilidad y comodidad de uso del escenario evaluado	170
Fig. 105 Opinión acerca de si esta plataforma soluciona un problema real.....	170
Fig. 106 Opinión sobre la mejora de la QoE percibida al utilizar la plataforma	170
Fig. 107 Opinión acerca del posible impacto en el mercado de la plataforma	171
Fig. 108 Resultados MOS \pm I.C.95% respecto a la experiencia de los participantes para cada dispositivo utilizado y respecto a la forma de consumir los contenidos relacionados.....	174
Fig. 109 Horas diarias utilizadas para consumir contenido online.....	174
Fig. 110 Horas diarias utilizadas para consumir contenido broadcast de TV	175
Fig. 111 Uso por parte de los participantes de diferentes dispositivos para el consumo de contenido online.....	175
Fig. 112 Experiencia previa de los participantes con el uso de dispositivos HMD	176
Fig. 113 Resultados del cuestionario SUS.....	178
Fig. 114 Inmersividad percibida por los participantes durante la evaluación	178
Fig. 115 Porcentaje de participantes que han perdido la noción del tiempo en algún momento	179
Fig. 116 Opinión respecto a la mejora que proporciona a la experiencia de consumo de TV la inclusión de contenidos omnidireccionales complementarios	179

Fig. 117 Temática o género del contenido para cuyo consumo los participantes utilizarían este escenario implementado en la plataforma desarrollada en la Tesis Doctoral.....	180
Fig. 118 Aspecto de la pantalla del terminal híbrido con el contenido visualizado en la MS app a pantalla completa (superior) o junto con la CS app de la herramienta de chat (inferior)	182
Fig. 119 CS app de la herramienta de chat en dispositivo secundario (smartphone).....	182
Fig. 120 Valor MOS e I.C. 95% obtenido en la segunda fase de la evaluación subjetiva.....	184
Fig. 121 Experiencia previa con escenarios o aplicaciones similares.....	185
Fig. 122 Problemas con plataformas similares	185
Fig. 123 Respuestas acerca del consumo compartido de contenido de TV en un mismo lugar	186
Fig. 124 Respuestas acerca del consumo compartido de contenido TV en diferentes lugares (TV Social).....	186
Fig. 125 Percepción de la distancia geográfica como impedimento para tener este tipo de experiencias compartidas	186
Fig. 126 Opinión respecto a cómo afecta a la QoE percibida por los usuarios este tipo de aplicaciones o escenarios.....	187
Fig. 127 Preferencias respecto a comunicarse a través de un chat privado sincronizado o público sin sincronizar (a través de redes sociales).....	188
Fig. 128 Opinión respecto a la utilidad de tener mensajes predefinidos contextualizados.....	188
Fig. 129 Percepción del nivel de sincronización alcanzado	188
Fig. 130 Percepción del funcionamiento del mecanismo de sincronización de la herramienta de chat	189
Fig. 131 Sensación de compañía durante la experiencia de uso del escenario	189
Fig. 132 Usabilidad del escenario.....	190
Fig. 133 Interés y utilidad del escenario	190
Fig. 134 Respuesta respecto a la utilidad real del escenario.....	190
Fig. 135 Posible impacto en el mercado actual según los participantes	191
Fig. 136 Uso de las aplicaciones si estuvieran disponibles	191
Fig. 137 Opinión sobre la capacidad de proporcionar experiencias de consumo más inmersivas, enriquecidas y personalizables	191
Fig. 138 Aplicabilidad de la plataforma en otras áreas.....	192
Fig. 139 Aspecto del escenario de evaluación subjetiva.....	193
Fig. 140 Experiencia previa con contenidos multisensoriales	195
Fig. 141 Opinión preliminar acerca de la mejora de la experiencia de consumo con la inclusión de efectos multisensoriales.....	195
Fig. 142 Percepción de sincronización entre contenidos presentados en la TV y los efectos multisensoriales.....	196
Fig. 143 Percepción de sincronización entre contenidos presentados en el HMD y los efectos multisensoriales.....	196
Fig. 144 Preferencia de los usuarios por algún efecto multisensorial	197
Fig. 145 Rechazo de los participantes hacia alguno de los efectos multisensoriales percibidos	197
Fig. 146 Opinión de los participantes acerca de la mejora de la QoE percibida si hay efectos multisensoriales presentes.....	197
Fig. 147 Interés de los participantes en poseer un sistema mulsemmedia (solo TV)	198
Fig. 148 Interés de los participantes en poseer un sistema mulsemmedia (TV + HMD)	198

Fig. 149 Predisposición de los participantes para ir a lugares o eventos donde se ofrezca experiencias de consumo mulsemia198

ÍNDICE DE TABLAS

Tabla 1 Clasificación de soluciones IDMS (adaptación de [9])	63
Tabla 2 Etiquetas y propiedades definidas en el fichero RMCF	79
Tabla 3 Conjunto de mensajes para el control de sesiones definidos para escenarios IDMS	87
Tabla 4 Ítems disponibles para el mensaje IDMS Join	88
Tabla 5 Conjunto de mensajes para el control de sincronización definidos para escenarios IDMS	89
Tabla 6 Tipos de acción del MS/CS app al MC	100
Tabla 7 Ejemplo de tabla de efectos mulsemmedia almacenada en la MS/CS app	101
Tabla 8 Etiqueta y propiedades integradas en el fichero RMCF para anunciar la disponibilidad de un servicio de chat	109
Tabla 9 Etiqueta y propiedades definidas para proporcionar un interfaz de mensajes contextualizados en la herramienta de chat	110
Tabla 10 Etiquetas del RMCF para soportar contenido omnidireccional	111
Tabla 11 Elementos GStreamer utilizados para la reproducción de contenido broadcast	116
Tabla 12 Elementos GStreamer utilizados para la reproducción de contenido tradicional broadband	120
Tabla 13 Contenidos híbridos utilizados en el escenario IDES básico	134
Tabla 14 Dispositivos utilizados en el escenario IDES básico	134
Tabla 15 Valor cuadrático medio e I.C. 95% del valor de asincronía obtenidos para el caso ideal en un terminal híbrido (ejecutando las MS y CS apps)	137
Tabla 16 Valor cuadrático medio e I.C. 95% de los valores de asincronía obtenidos para el caso ideal en entorno multidispositivo (apps MS y CS ejecutándose en dispositivos diferentes)	138
Tabla 17 Valor cuadrático medio e I.C. 95% de los valores de las asincronías registradas en un escenario realista utilizando un único terminal híbrido (ejecutando las MS y CS apps)	139
Tabla 18 Valor cuadrático medio e I.C. 95% del valor de las asincronías registradas en un escenario realista multidispositivo (MS y CS apps ejecutándose en dispositivos diferentes)	139
Tabla 19 Dispositivos involucrados en el escenario IDES inmersivo	142
Tabla 20 Características del contenido omnidireccional	142
Tabla 21 Características del contenido generado para la evaluación	144
Tabla 22 Valor cuadrático medio (RMS) e I.C. 95% de los valores de asincronía registradas	144
Tabla 23 Dispositivos involucrados en la evaluación del escenario IDMS	148
Tabla 24 Resumen de los casos considerados y tipo(s) de sincronización evaluado(s) en cada uno de ellos	151
Tabla 25 Resultados obtenidos del RMS medio e I.C. 95% de los valores de la asincronía registrados en las 10 sesiones realizadas en cada caso respecto a la solución de sincronización IDMS con un esquema de control SMS	152
Tabla 26 Resultados obtenidos del RMS medio e I.C. 95% de los valores de la asincronía registrados en las 10 sesiones realizadas en cada caso respecto a la solución de sincronización IDMS con un esquema de control DCS	153
Tabla 27 Valor cuadrático medio e I.C. 95% de los valores de las asincronías registradas con un esquema de control centralizado SMS	155
Tabla 28 Valor cuadrático medio e I.C. 95% de los valores de las asincronías registradas con un esquema de control distribuido DCS	155
Tabla 29 Probabilidad de registrar un valor de asincronía (valor absoluto) de hasta 80ms	157

Tabla 30 Contenidos híbridos utilizados en el escenario IDES básico	164
Tabla 31 Dispositivos utilizados en el escenario IDES básico	165
Tabla 32 Perfil de los participantes	165
Tabla 33 Casos evaluados en el primer test	166
Tabla 34 Escala Likert utilizada para calcular la métrica MOS.....	166
Tabla 35 Contenido generado para la evaluación subjetiva	172
Tabla 36 Perfil de los participantes	172
Tabla 37 Cuestionario SUS.....	177
Tabla 38 Dispositivos involucrados en la evaluación del escenario IDMS.....	181
Tabla 39 Perfil de los participantes	183
Tabla 40 Contenido generado para la evaluación subjetiva	192
Tabla 41 Efectos multisensoriales incluidos en el escenario de evaluación	193
Tabla 42 Perfil de los participantes	194

ÍNDICE DE ECUACIONES

Ecuación 1 Cálculo de la duración del ajuste adaptativo del proceso de reproducción AMP	81
Ecuación 2 Cálculo del instante de generación objetivo	83
Ecuación 3 Cálculo del nivel de sincronización alcanzado por una CS app en un escenario mono- o multidispositivo	84
Ecuación 4 Cálculo de la ronda de sincronización.....	92
Ecuación 5 Cálculo de la referencia de sincronización IDMS a un punto medio del proceso de reproducción	97
Ecuación 6 Cálculo del instante de activación de los efectos multisensoriales.....	102

ACRÓNIMOS

En la presente memoria, se emplean los siguientes acrónimos:

AF	Adaptation field
AIT	Application Information Table
AMP	Adaptive Media Playout
ATS	Advanced Transport Scheme
ATSC	Advanced Television Systems Committee
BER	Bit Error Rate
CBR	Constant Bit Rate
CFD	Cummulative Frequency Distribution
CMP	Cube Map based Projection
CS	Companion Screen
CTS	Composition Timestamp
DASH	Dynamic Adaptive Streaming over HTTP
DASH-SRD	DASH-Spatial Representation Description
DCS	Distributed Control Scheme
DLNA	Digital Living Network Alliance
DNS	Domain Name Server
DNS-SD	DNS-Service Discovery
DoF	Degrees of Freedom
DTS	Decoding Timestamp
DTT	Digital Terrestrial TV
DVB	Digital Video Broadcasting
DVB-CSS	DVB Companion Screens and Streams
DVB-T	Digital Video Broadcasting – Terrestrial
EIT	Event Information Table
EPG	Electronic Program Guide
ERP	Equi Rectangular based Projection
ES	Elementary Stream
ETSI	European Telecommunications Standards Institute
FoV	Field of View
GOP	Group Of Pictures
GUI	Graphical User Interface
HAS	HTTP Adaptive Streaming
HBBTV	Hybrid Broadcast Broadband TV
HCI	Human Computer Interaction
HEVC	High Efficiency Video Coding
HLS	HTTP Live Streaming
HMD	Head Mounted Display
HTTP	Hypertext Transfer Protocol
IDES	Inter Device Media Synchronization
IDMS	Inter Destination Media Synchronization
IEC	International Electrotechnical Commission
IP	Internet Protocol
IPTV	Internet Protocol Television
ISO	International Organization for Standardization
ITU	International Telecommunication Union
LAN	Local Area Network
MHP	Multimedia Home Platform
MMT	MPEG Media Transport

MPD	Media Presentation Description
MPEG	Moving Pictures Experts Group
MPEG2-TS	MPEG2-Transport Stream
M/S	Master Slave Scheme
MS	Main Screen
NAT	Network Address Translation
NCL	Nested Content Language
NTP	Network Time Protocol
OTT	Over The Top
PAL	Phase Alternating Line
PCR	Program Clock Reference
PES	Packetized Elementary Stream
PIP	Picture in Picture
PMT	Program Mapping Table
PTP	Precision Time Protocol
PTS	Presentation Timestamp
QEF	Quasi Error Free
QoE	Quality of Experience
QoS	Quality of Service
RDS	Radio Data System
RFC	Request For Comments
RTCP	RTP Control Protocol
RTP	Real-Time Transport Protocol
RTT	Round Trip Time
RV	Realidad Virtual
SEDL	Sensory Effect Description Language
SEM	Sensory Effect Metadata
SEV	Sensory Effect Vocabulary
SMS	Synchronization Maestro Scheme
SO	Sistema Operativo
STB	Set Top Box
TDT	Time and Date Table
TEMI	Timing and External Media Information
TCP	Transmission Control Protocol
T.D.	Tesis Doctoral
UDP	User Datagram Protocol
UHD	Ultra High Definition
UHDTV	UHD Television
UPnP	Universal Plug and Play
URL	Uniform Resource Locator
URI	Uniform Resource Identifier
UTC	Coordinated Universal Time
VOD	Video On Demand – Vídeo bajo demanda
WC	Wall Clock
WLAN	Wireless LAN – Red de área local inalámbrica
WS	Websocket
WSS	WebSocket Server

Capítulo 1. INTRODUCCIÓN Y OBJETIVOS DE LA TESIS DOCTORAL

En este capítulo se expone una introducción al campo en el que se enmarca la Tesis Doctoral, así como sus objetivos y subobjetivos. Además, también se describe la estructura de esta memoria.

1.1 INTRODUCCIÓN Y MARCO DE LA TESIS DOCTORAL

Tradicionalmente, los servicios multimedia estaban concebidos para ser transmitidos a través de redes específicas y consumidos en dispositivos concretos. Un claro ejemplo de esto es la televisión tradicional, cuyos contenidos se transmiten a través de una red broadcast (*Digital Video Broadcasting* - DVB) y se consumen a través de un televisor.

Hoy en día, existe una gran variedad de tecnologías de entrega de contenidos, dispositivos de consumo y contenidos multimedia a disposición de los consumidores. Respecto a la entrega de contenidos multimedia, ésta puede realizarse a través de tecnologías broadcast o broadband. Por un lado, las tecnologías broadcast (por ejemplo, DVB) permiten transmitir un mismo contenido a un gran número de consumidores de manera simultánea. Por otro lado, las tecnologías broadband (por ejemplo, IP) proporcionan servicios bidireccionales, interactivos y adaptativos, según los recursos disponibles o las preferencias de los consumidores. Sin embargo, las tecnologías broadband proporcionan un peor rendimiento en factores como la escalabilidad, estabilidad o latencia, comparadas con las tecnologías broadcast. Además, ofrecen diferentes métodos para la entrega de contenidos, como son la descarga y el streaming.

Existe una amplia variedad de tecnologías de streaming de contenido, según si el servicio a ofrecer es gestionado o no gestionado [1]. Por un lado, los servicios gestionados (como, por ejemplo el servicio IPTV) se ofrecen en entornos controlados y normalmente se basan en el uso de protocolos como Real-Time Transport Protocol junto con RTP Control Protocol (RTP y RTCP, respectivamente), definidos en la RFC 3550 [2]. De esta manera, la transmisión del contenido puede ser *push-based*, es decir, es el transmisor quien ordena el envío de los flujos del contenido. Por otro lado, los servicios no gestionados (como, por ejemplo, la TV a través de la Web o WebTV) pueden ofrecerse en redes no controladas (como, por ejemplo, Internet), y así operar globalmente. Se suelen basar en tecnologías *pull-based* (es decir, es el consumidor quien demanda los flujos de contenido) y adaptativas (por ejemplo, en técnicas de streaming adaptativo basado en HTTP, *Http-based Adaptive Streaming* o HAS). Las principales ventajas de los servicios no gestionados frente a los gestionados son la escalabilidad, adaptabilidad, fiabilidad, ubicuidad y eficiencia en términos de costes. Hoy en día, las técnicas HAS están siendo adoptadas de manera general para la transmisión de contenido en redes broadband. De hecho, existen diferentes soluciones y estándares de soluciones de HAS, como *Http Live Streaming* (HLS), de Apple [3]; *Dynamic Adaptive Streaming over Http* (DASH), de ISO/IEC MPEG [4]; *Http Dynamic Streaming*

(HDS), de Adobe¹; o *Microsoft Smooth Streaming* (MSS), de Microsoft². Cabe destacar que la solución DASH ha sido adoptada por especificaciones como HbbTV [5], y por proveedores de contenidos multimedia, como, por ejemplo, Youtube o Netflix (según el Sistema Operativo o SO sobre el que se ejecute).

Respecto a los dispositivos de consumo multimedia, sin tener en cuenta la amplia variedad de TVs conectadas, la proliferación y utilización de diferentes dispositivos secundarios para dicho fin es ya una realidad. Por una parte, estos dispositivos cuentan con una compatibilidad para múltiples tecnologías de conexión (por ejemplo, WiFi, 3G, 4G, 5G...), por lo que son capaces de recibir contenidos multimedia a través de una (o varias) tecnologías de transmisión. Por otra parte, estos dispositivos secundarios cuentan con una gran heterogeneidad de recursos, rendimiento y capacidades, sobre todo en términos de disponibilidad de ancho de banda, soporte a diferentes tipos de tecnología o a tipos de contenido multimedia, capacidad de procesamiento y tamaño de las pantallas.

Este ecosistema tan variado (y variable) puede inducir a una rivalidad entre tecnologías y fabricantes de dispositivos que termine generando problemas de compatibilidad generalizados. Sin embargo, también proporciona una oportunidad única para la investigación y el desarrollo, cuyo objetivo principal sea la convergencia, coordinación e interoperabilidad entre las diferentes tecnologías de transmisión existentes. De esta forma, dos tecnologías (broadcast y broadband) que, a priori, se concibieron para ser empleadas de manera independiente, pueden combinarse para complementarse y proporcionar al consumidor un único ecosistema híbrido para un consumo enriquecido de contenido multimedia. Además, otro de los objetivos esenciales para aumentar y mejorar la experiencia de consumo multimedia es conseguir proporcionar una interacción y coordinación, transparentes para el usuario, entre los dispositivos de consumo involucrados. En particular, esto debería ser relevante para operadores de TV y otras empresas relacionadas interesadas (por ejemplo, fabricantes de dispositivos, proveedores de contenidos multimedia, productoras, etc.) puesto que el contenido broadcast (el consumido en la TV) puede ser ampliado y complementado por otro contenido relacionado (en directo o bajo demanda), que se pueda descargar (por streaming) a través de redes broadband (por ejemplo, Internet), enriqueciendo así los servicios multimedia ofrecidos.

Tal y como se muestra en [6], existen una serie de escenarios o casos relevantes en los que los servicios multimedia híbridos mejoran y enriquecen de forma evidente la experiencia de consumo de los usuarios. Alguno de estos escenarios o casos son la provisión de diferentes flujos de audio a través de la red broadband, para permitir la sustitución del audio transmitido vía broadcast; la oferta de contenido de TV multidispositivo y la disponibilidad de selección multivista; el consumo simultáneo de dos o más contenidos multimedia en uno o diferentes dispositivos (por ejemplo, mediante vistas solapadas, *Picture in Picture* o PiP, o en mosaico, si se utiliza un solo dispositivo); disponibilidad de escalabilidad temporal, espacial o de color [7]; *tiled-streaming* (esto es, contenido multimedia en el que las regiones de la imagen en el flujo de vídeo se transmiten de manera independiente a la hora de solicitar una determinada calidad a través de flujos diferentes) [8]; etc.

Este nuevo escenario de consumo enriquecido de contenidos relacionados con la TV y los nuevos servicios híbridos emergentes, pueden generar un gran impacto en el sector del entretenimiento. Además, también se pueden proporcionar beneficios sociales, de manera que los usuarios pueden sentirse más integrados durante la experiencia de consumo multimedia. Algunos ejemplos serían la posibilidad de seleccionar y consumir de forma simultánea al

¹ <http://www.adobe.com/products/hds-dynamic-streaming.html>

² <https://www.microsoft.com/silverlight/smoothstreaming/>

contenido principal de la TV contenido broadband con el audio nativo; subtítulos configurables o audio descripción; o incluso la visualización de un vídeo complementario incluyendo a un intérprete de lenguaje de signos. Algunas de estas características pueden resultar muy valiosas para usuarios de edad avanzada o con algún tipo de discapacidad.

Proporcionar todo este tipo de servicios empleando únicamente tecnología broadcast sería técnicamente muy costoso y, por tanto, improbable. Sin embargo, el ancho de banda adicional necesario (de alto coste y escaso en el dominio broadcast) y el hecho de que no todos los consumidores estarían dispuestos a pagar por recibir contenidos complementarios, son algunos de sus inconvenientes. Por tanto, adoptar la tecnología de transmisión broadband para estas funciones se convierte en una estrategia más eficaz, puesto que proporciona flexibilidad en la demanda particular de contenidos y permite la configuración y adopción de mecanismos adaptativos, que permiten que dispositivos con una gran heterogeneidad de especificaciones técnicas, puedan utilizarse para consumir los contenidos multimedia involucrados en el servicio híbrido.

Por tanto, se puede observar que se está produciendo una evolución con relación a la forma de consumir contenido, y dicha evolución permite mejorar la calidad de la experiencia (*Quality of Experience* o QoE) del consumidor, aumentando la disponibilidad de dispositivos de consumo y proporcionando unos servicios multimedia más adaptativos y omnipresentes.

Adicionalmente, el consumo en grupo de contenido audiovisual ha sido un hábito social (por ejemplo, una familia en el salón de su hogar viendo la TV). Sin embargo, el consumo compartido de contenidos multimedia no siempre es posible en la sociedad actual, donde familiares, amigos o compañeros no tienen por qué vivir en la misma casa, ciudad o región. En ese contexto, la aparición de nuevas tecnologías, junto con el uso de redes sociales y sistemas de audio/videoconferencia, permiten recrear dichos escenarios entre usuarios que están geográficamente separados.

Con el fin de poder proporcionar una experiencia de consumo satisfactoria entre usuarios remotos, aún se deben afrontar una serie de retos tecnológicos. Como ejemplo, algunos de ellos son la escalabilidad, la integración de redes sociales, la privacidad, la Calidad de Servicio y/o Experiencia (*Quality of Service/Experience*, o QoS/E) o la sincronización multimedia. Concretamente, esta Tesis Doctoral se centra en buscar soluciones al reto de la sincronización multimedia en dichos tipos de escenarios.

Uno de los casos de uso identificados en [9], en los que se requiere de mecanismos de sincronización para proporcionar una calidad de experiencia (QoE) satisfactoria a los consumidores, es el caso de *TV Social*. Este caso de uso permite que diferentes grupos de usuarios ubicados en diferentes zonas geográficas, e independientemente de la tecnología de entrega de contenidos utilizada, consuman los contenidos de un programa de TV de manera simultánea, pudiendo, de esta manera, interactuar y compartir servicios, a través de herramientas como la de chat o videoconferencia. El caso de uso de *TV Social* combina contenido de TV con la interacción a través de redes sociales entre usuarios.

Por un lado, cuando usuarios remotos están disfrutando de una experiencia compartida de consumo de TV, mecanismos de sincronización entre destinatarios (*Inter-Destination Media Synchronization* o IDMS) se vuelven imprescindibles para poder ofrecer una experiencia satisfactoria. Dichos mecanismos son los encargados de que los usuarios remotos, que comparten la experiencia, consuman, de forma simultánea, los mismos instantes del contenido mientras interactúan entre sí. De otro modo, si no existieran estos mecanismos, la QoE percibida por alguno de los usuarios remotos involucrados podría verse perjudicada en determinados instantes. Como ejemplo, en un escenario de consumo de un evento deportivo entre un grupo de usuarios separados geográficamente, que cuentan con un canal de chat para interactuar, no es aceptable

que alguno de los usuarios comente alguna jugada o acción que algún usuario aún no haya visualizado (por ejemplo, un gol en un partido de fútbol que un usuario esté comentando en un canal de chat mientras hay otros usuarios del mismo grupo que no han visto todavía cómo se dispara el balón a portería).

Para alcanzar un estado de sincronización (IDMS) satisfactorio, es necesario compensar los diferentes retardos que puedan existir en las redes de distribución involucradas (broadcast y broadband, si la TV recibe la señal a través de redes de radiodifusión, como, por ejemplo, DVB; o únicamente broadband, si la TV recibe la señal a través de redes de banda ancha, como, por ejemplo, IPTV). Los retardos variables pueden provocar que la entrega de los contenidos no sea simultánea, es decir, que no se entreguen a los diferentes usuarios de manera sincronizada sin establecer ningún tipo de mecanismo de ajuste posterior (es decir, sin la necesidad de adoptar técnicas de IDMS). Esto puede conllevar a que la experiencia compartida de consumo de TV provoque interacciones incoherentes que desemboquen en frustración por parte de algún usuario final.

Por otro lado, actualmente existe un auge de consumo de contenido accesible vía broadband (es decir, Internet) relacionado (y/o complementario) al retransmitido vía broadcast (por ejemplo, DVB). De hecho, el auge de la disponibilidad de contenido a través de Internet ha permitido ofrecer al consumidor un gran abanico de diferentes tipos de contenidos que requieren un mayor ancho de banda, como contenido omnidireccional (también conocido como *contenido 360*) o contenido UHD (*Ultra High Definition*) o 4K, cuyo envío a través de redes broadcast supone un coste significativamente más alto. Una evidencia del crecimiento de interés por este tipo de ecosistema, donde se proporcionan contenidos a través de redes híbridas, es el reciente estándar *Hybrid Broadcast Broadband TV* (HbbTV) [5]. Dicho estándar pone las bases para poder proporcionar experiencias interactivas a los usuarios, mediante la combinación de contenido recibido por redes broadcast (en adelante, *contenido broadcast*) con contenidos recibidos por redes broadband (en adelante, *contenido broadband*). De esta forma, se pueden obtener muchas ventajas, como, por ejemplo: una mayor interactividad con el contenido, información segmentada al perfil del usuario y, a partir de su versión 2.0, la utilización de dispositivos secundarios (por ejemplo, smartphones o tablets) como pantallas secundarias, en las que poder consumir contenido multimedia broadband relacionado con el contenido broadcast que está siendo visualizado en el TV (es decir, el dispositivo principal). De hecho, el consumo de contenido broadband a través de dispositivos secundarios está incrementando día a día. Como evidencia de ello, según un estudio del año 2020 [10], el 90% de usuarios de Internet afirman consumir contenido multimedia online, de los cuales un 53.3% lo hace a través de este tipo de dispositivos. En particular, respecto al estándar HbbTV, se puede afirmar que el potencial que puede llegar a ofrecer está aún por llegar, puesto que en una encuesta en [11], más de la mitad de los encuestados veían como atractivo (45.2%) o muy atractivo (23.2%) poder disponer de un entorno de consumo híbrido, a pesar de que un 35.8% de los encuestados afirmaron tener la sensación de que el despliegue de este tipo de servicio iba por detrás de las expectativas ofrecidas.

Resulta evidente que, cuando dos o más dispositivos están involucrados en la experiencia de consumo, también es necesario establecer mecanismos de sincronización entre los procesos de reproducción de todos ellos (mecanismos conocidos como *Inter-Device Media Synchronization* o IDMS) con el objetivo de proporcionar una experiencia de consumo satisfactoria. Para este tipo de escenarios, si los contenidos multimedia consumidos se reciben a través de tecnologías de distribución diferentes (broadcast y/o broadband), la sincronización requerida pasa a denominarse **sincronización híbrida**. Por tanto, en entornos inter-destinatario e inter-dispositivo con contenidos híbridos, serán necesarios mecanismos de sincronización híbrida tanto IDMS como IDMS.

Respecto al contenido disponible a través de tecnología broadband, es evidente la relevancia de los servicios OTT (*Over the Top*) de consumo de contenido bajo demanda (*Video on Demand* o VoD) a través de aplicaciones como, por ejemplo, Netflix, HBO o Amazon Prime. Además, recientemente, el vídeo omnidireccional (también llamado vídeo 360) también ha experimentado una mayor demanda. Tal y como se anunciaba en [12], en 2019 ya existía un número creciente de usuarios que ya poseían un visor de vídeo 360 (*Head Mounted Display* o HMD) en el hogar, por lo que ya se preveía que este tipo de dispositivos iban a ser habituales en los hogares a corto/medio plazo y, por tanto, la variedad y la oferta de contenidos omnidireccionales específicos también sería mayor. De hecho, la Radio Televisión Pública Española (RTVE) ya ha ofrecido, de manera experimental, algún tipo de contenido omnidireccional en series como "El Ministerio del Tiempo" o "El Caso".

A pesar del incremento de consumo de contenido a través de redes broadband, el consumo de TV a través de redes broadcast sigue siendo el medio de entretenimiento más utilizado [13], [14]. Aunque el contenido a través de tecnología broadband está teniendo un impacto real, ofreciendo a los usuarios una interactividad y flexibilidad que la tecnología broadcast no proporciona, en Europa, la Televisión Digital Terrestre o TDT (basada en DVB-T) continúa creciendo de manera constante³, por lo que la evolución hacia una televisión híbrida que combine las ventajas de ambas tecnologías de transmisión resulta prometedora. El éxito y la aceptación por parte de los consumidores del modelo híbrido demuestra que el consumo de contenido multimedia exclusivamente a través de la red broadband (por ejemplo, IPTV) sigue siendo un escenario improbable en los próximos años⁴. De hecho, según el estudio presentado en [15], en España, en 2019 había un 36% de ciudadanos mayores de 16 años que no consumían contenido digital y, entre ellos, una tercera parte ni siquiera conocían la oferta de contenido adicional o complementario al contenido broadcast a través de redes broadband (esto es, no conocían la existencia de contenidos adicionales a través de Internet desde su TV). En ese mismo estudio, sin embargo, se afirmó que el 71% de los encuestados que consumían contenido multimedia a través de Internet, lo hacían utilizando una TV conectada. Concretamente, un 5.6% utilizaban únicamente una TV conectada y el restante 65.4% utilizaban, además, dispositivos como smartphones, set top boxes, tablets, consolas, etc.

Así como la oferta de contenido a través de redes broadband ha experimentado un gran auge, recientemente la experiencia de consumo de contenido ha dejado de involucrar exclusivamente a los sentidos de la vista y el oído. Con el fin de estimular el resto de los sentidos, surge un nuevo concepto de contenido denominado *mulsemedia* (de *multisensorial media*). El contenido *mulsemedia* permite proporcionar al usuario estímulos para otros sentidos, además de los dos citados, a través de otro tipo de dispositivos adicionales que permiten la generación de efectos sensoriales, como, por ejemplo, viento, calor, frío, luz ambiental, niebla, vibración, aromas, etc. En definitiva, el principal objetivo de incluir este nuevo tipo de contenido en las experiencias de consumo multimedia es proporcionar al usuario una experiencia más inmersiva y realista ("*Ver es creer, pero sentir es la realidad*" -T. Fuller [16]).

Por tanto, se puede afirmar que el nuevo paradigma de experiencia de consumo multimedia que se presenta cuenta con las tecnologías y dispositivos necesarios para poder ser social, inmersivo y multisensorial. Por lo tanto, para asegurar una buena QoE a los usuarios, resulta necesario, entre otros desarrollos complementarios, el diseño y establecimiento de mecanismos de sincronización híbridos que sean agnósticos e independientes de la tecnología de distribución de los contenidos.

³ <https://dtg.org.uk/the-dtg-summit-2018/>

⁴ <http://www.feh-mi.com/blog/the-dtt-platform-succeeding-in-the-faceof-the-svod-challenge/>

En esta Tesis Doctoral, se estudian los diferentes mecanismos de sincronización híbridos, y centra su foco en el caso de uso de *TV Social*, diseñando una solución de sincronización híbrida para escenarios IDMS e IDES, generando contenido multisensorial y validando dicha solución en una plataforma híbrida compatible con el estándar HbbTV. Se aborda, principalmente, el diseño, implementación y validación de mecanismos de sincronización híbridos, teniendo en cuenta la naturaleza del contenido (multimedia y/o mulsemedia) y su entorno de consumo (multidispositivo y/o multidestinario), validando posteriormente la calidad de la experiencia percibida por parte de los consumidores (QoE).

1.2 OBJETIVOS DE LA TESIS DOCTORAL

El **objetivo principal** de la Tesis Doctoral es *diseñar, implementar y evaluar diferentes soluciones para facilitar la reproducción de contenidos de TV híbridos (broadcast/broadband) y efectos multisensoriales relacionados, de forma sincronizada, mediante el uso de uno o varios dispositivos de consumo, tanto en un entorno local (soluciones IDES) como en un entorno distribuido (soluciones IDMS).*

Para conseguir dicho objetivo principal, se deberán conseguir los siguientes **sub-objetivos**:

- I. Estudiar y analizar las diferentes propuestas existentes para combinar y relacionar contenidos de TV híbridos y multisensoriales, con la finalidad de identificar sus ventajas y desventajas.
- II. Estudiar las diferentes tecnologías actuales de distribución broadcast y broadband de contenidos multimedia. Analizar las limitaciones y posibilidades para conseguir la sincronización híbrida.
- III. Estudiar los diferentes tipos de contenidos susceptibles de ser incluidos en sistemas multimedia multisensoriales híbridos, como, por ejemplo, de tipo textual, de aplicaciones de Social Media, de tipo sensorial, así como de contenido multimedia tradicional y 360.
- IV. Identificar el problema de la heterogeneidad a nivel de fuentes y tipos de contenido, formatos de codificación, redes y tecnologías de transmisión, así como de dispositivos de consumo.
- V. Realizar un estudio del estado del arte relacionado con la sincronización de contenidos híbridos.
- VI. Realizar un estudio para obtener información sobre preferencias, hábitos y expectativas de los usuarios.
- VII. Definir y especificar soluciones de sincronización híbrida, IDES e IDMS, compatibles con las tecnologías de distribución más extendidas, tanto de la reproducción de contenidos como también de la generación de efectos multisensoriales asociados a los contenidos audiovisuales.
- VIII. Crear una plataforma extremo-a-extremo para la generación, distribución (adaptativa) y consumo de contenidos híbridos relacionados, en la que poder implementar prototipos de aplicaciones reales con contenidos híbridos y efectos multisensoriales, utilizando software de código abierto y modificándolo según las necesidades de la Tesis Doctoral. Esta plataforma debe incluir el diseño, implementación e integración en la misma de dispositivos de generación de efectos multisensoriales comerciales o de

diseño propio. También debe incluir una solución para poder enlazar contenidos híbridos relacionados, así como para señalar, publicitar y actualizar su existencia y facilitar su selección por parte de los dispositivos de consumo. Adicionalmente, para poder incluir efectos multisensoriales en la plataforma, será necesario definir métodos adecuados para codificar, etiquetar, insertar y transmitir información sobre dichos efectos.

- IX. Diseñar e implementar en la plataforma varios prototipos que incluyan contenidos de TV híbridos, efectos multisensoriales y las soluciones de sincronización desarrolladas.
- X. Evaluar objetiva y subjetivamente las soluciones de sincronización propuestas en los diferentes prototipos implementados.

1.3 METODOLOGÍA DE LA TESIS DOCTORAL

Para el desarrollo de esta Tesis Doctoral, el proceso seguido se puede resumir en las siguientes etapas:

Etapa 1: Realización de un análisis exhaustivo del estado del arte, estudiando los diferentes tipos de contenido multimedia, efectos multisensoriales y sus métodos de generación y de señalización, las tecnologías de distribución más comunes en comunicaciones híbridas y los mecanismos de sincronización propuestos hasta la fecha. En esta etapa se identificaron las limitaciones y ventajas de cada una de las soluciones existentes.

Etapa 2: Recopilación de casos de uso en los que se requiere sincronización híbrida IDES e IDMS, así como la inclusión de efectos multisensoriales. En esta etapa se identificaron los requisitos necesarios para proporcionar una experiencia de consumo satisfactoria en el consumidor final (como, por ejemplo, alcanzar un nivel de sincronización entre los procesos de reproducción de contenidos relacionados imperceptible para el consumidor), así como los retos asociados a dichos requisitos, con tal de abordarlos en las etapas posteriores.

Etapa 3: En paralelo, en el grupo I+D en el que se enmarca la presente Tesis Doctoral, se realizó un estudio de mercado orientado a detectar las posibles aplicaciones de los servicios híbridos en cuanto a la producción y distribución de contenidos. En esta etapa se realizaron una serie de encuestas a más de 1000 usuarios en España y se obtuvo información acerca de sus hábitos en cuanto al consumo de contenidos multimedia, su visión/opinión acerca de los beneficios que se pueden aportar, así como sus preferencias y expectativas. En base a los casos de uso propuestos (como por ejemplo, la *TV Social*, o el escenario de consumo de contenido multipantalla), requisitos (como, por ejemplo, un nivel de sincronización que haga imperceptible las diferencias temporales entre los procesos de reproducción de contenidos relacionados) y premisas de diseño (como, por ejemplo, que sea compatible con el estándar HbbTV), se abordó y documentó la especificación de los diferentes componentes de sincronización necesarios.

Etapa 4: Desarrollo de soluciones de sincronización híbrida IDES, IDMS y de efectos multisensoriales.

Etapa 5: Definición de un método de señalización de efectos multisensoriales.

Etapa 6: Desarrollo de la plataforma extremo-a-extremo para la generación, distribución y consumo sincronizado de contenidos híbridos, incluyendo contenidos omnidireccionales y efectos multisensoriales. Dicha plataforma se ha desarrollado empleando librerías open-source como GStreamer y los lenguajes de programación C++, Java y HTML5, JS, CSS3; tal y como se describe en los capítulos correspondientes.

Etapa 7: Implementación en la plataforma de una serie de prototipos de caso de uso y evaluación de las soluciones presentadas en la Tesis Doctoral, tanto de forma objetiva (esto es, la precisión de la sincronización alcanzada) como de forma subjetiva (percepción del nivel de sincronización, experiencia de uso, nivel de mejora conseguida mediante la inclusión de efectos multisensoriales en los prototipos, etc.). Los resultados obtenidos han servido para ir mejorando los prototipos finales.

Etapa 8: En paralelo, publicación de los resultados parciales y globales que se hayan ido obteniendo.

1.4 ORGANIZACIÓN DE LA MEMORIA

La memoria se organiza de la manera descrita a continuación. En el Capítulo 2 se hace una revisión de los conceptos, estándares y protocolos que permiten el establecimiento de mecanismos de sincronización para un consumo interactivo e inmersivo de contenidos multimedia y mulsemmedia.

En el Capítulo 3, se exponen los principales trabajos relacionados con la Tesis Doctoral, en áreas como la sincronización híbrida, TV Social, y mulsemmedia.

En el Capítulo 4, se presentan las soluciones de sincronización multimedia y multisensorial híbrida, IDES e IDMS diseñadas. Además, se incluyen las tecnologías adoptadas en el desarrollo de esta Tesis Doctoral

En el Capítulo 5, se describe el diseño y desarrollo de la plataforma extremo-a-extremo de consumo híbrido multimedia, multidestinatarario y multidispositivo, incluyendo los mecanismos de sincronización propuestos para cada uno de dichos escenarios, así como efectos multisensoriales.

En el Capítulo 6, se presenta la metodología y los resultados obtenidos en la evaluación objetiva respecto a los niveles de sincronización alcanzados para cada tipo de requerimiento dentro de la plataforma desarrollada. Por un lado, se presenta el nivel de sincronización alcanzado entre dispositivos dentro de la misma red de área local (IDES), es decir, entre un dispositivo principal y otros dispositivos secundarios, como segundas pantallas. Por otro lado, se presenta el nivel de sincronización alcanzado entre dispositivos en entornos distribuidos, entre destinatarios remotos (IDMS).

En el Capítulo 7, se presenta la metodología seguida y los resultados obtenidos en la evaluación subjetiva llevada a cabo. En ella se ha recogido, entre otros parámetros relevantes, las preferencias, expectativas y el nivel de QoE percibida por los consumidores a través de la plataforma extremo-a-extremo desarrollada. En los prototipos implementados en la plataforma se han evaluado los diferentes tipos de mecanismos de sincronización implementados de manera independiente y, también, de manera combinada.

Finalmente, en el Capítulo 8, se presentan las conclusiones de la Tesis Doctoral, así como un resumen del trabajo futuro a realizar, como posible continuación de la misma.

Capítulo 2. CONCEPTOS Y MECANISMOS RELACIONADOS

En este capítulo se presentan los conceptos, componentes y tecnologías clave, así como una revisión de trabajos sobre estándares y protocolos que permiten la distribución, recepción, y consumo sincronizado, interactivo e inmersivo de contenido multimedia y multisensorial.

2.1 TIPOS DE SINCRONIZACIÓN MULTIMEDIA

A continuación, se definen los diferentes tipos de sincronización multimedia, puesto que son conceptos que se van a emplear a lo largo de este documento:

- Los mecanismos de sincronización **intraflujo** se encargan del mantenimiento, durante la reproducción del contenido, de la relación temporal entre las unidades de dato de cada flujo (por ejemplo, que los fotogramas de una secuencia de vídeo se muestren en orden y durante el tiempo apropiado).
- Los mecanismos de sincronización **interflujo** se encargan de que los diferentes flujos en un contenido multimedia se presenten alineados temporalmente entre ellos (por ejemplo, que el contenido de audio corresponda con el contenido vídeo que se está visualizando).
- Los mecanismos de sincronización **interdispositivo** (*Inter Device Synchronization, IDES*) se adoptan en escenarios en los que hay más de un dispositivo de consumo y cercanos (por ejemplo, en una LAN) dichos mecanismos permiten que estos dispositivos reproduzcan contenidos multimedia relacionados de forma sincronizada.
- Los mecanismos de sincronización **interdestinatario** (*Inter Destination Media Synchronization, IDMS*) también se adoptan cuando hay más de un dispositivo de consumo pero que no están próximos unos a otros (por ejemplo, dispersos geográficamente). Se adoptan para alcanzar un nivel de sincronización satisfactorio entre dispositivos separados geográficamente.

2.2 ENTREGA HÍBRIDA DE CONTENIDOS

En un sistema de entrega de contenidos híbrido, M flujos multimedia pueden ser generados por S fuentes, y entregados a D dispositivos a través de N redes empleando, para tal fin, hasta T tecnologías de transmisión diferentes. En el contexto de esta Tesis Doctoral, la entrega de contenidos se considera híbrida cuando $N > 1$ y/o $T > 1$. En la Fig. 1, puede observarse el escenario más simple en el que $M = S = N = T = n$; siendo $n > 1$, con n fuentes generando n contenidos relacionados siendo entregados en n flujos a través de n redes heterogéneas (híbridas) a un único (líneas discontinuas) o a n dispositivos de consumo. Adicionalmente, se denomina terminal híbrido a aquel dispositivo de consumo capaz de recibir e interpretar contenido transmitido a través de $T > 1$ tecnologías de transmisión (por ejemplo, el Dispositivo D1 en la Fig. 1).

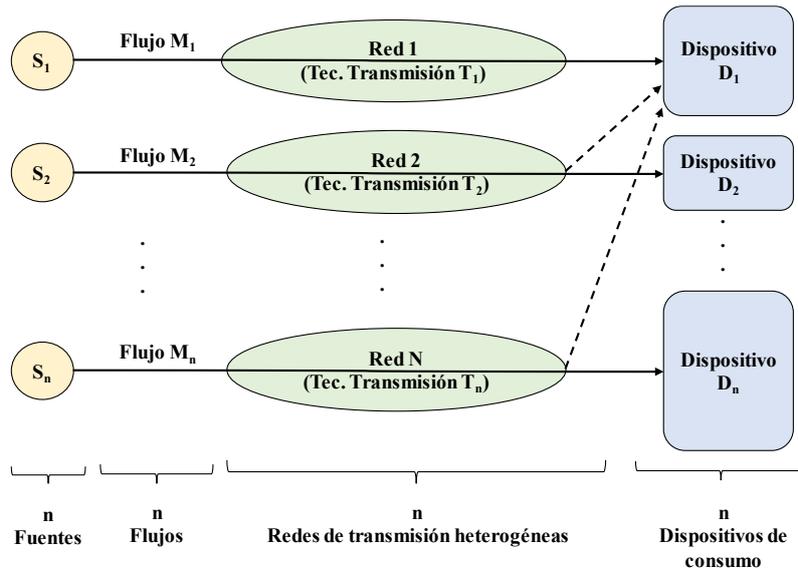


Fig. 1 Escenario de entrega de contenidos multimedia híbridos

2.3 PROBLEMA DE LA VARIABILIDAD DE RETARDOS EN ESCENARIOS HÍBRIDOS.

Uno de los principales problemas que presenta la entrega y consumo de contenido híbrido es la alta variabilidad que presentan los diferentes retardos de red entre las fuentes de los contenidos y los dispositivos de consumo de los mismos. En este tipo de escenarios, las cadenas de transmisión involucradas pueden introducir una serie de artefactos temporales (latencias) que provocan diferentes retardos entre flujos y entre destinatarios. Como ejemplo, en la Fig. 2 puede observarse la diferencia entre retardos extremo-a-extremo de diferentes flujos multimedia que han sido generados en el mismo instante.

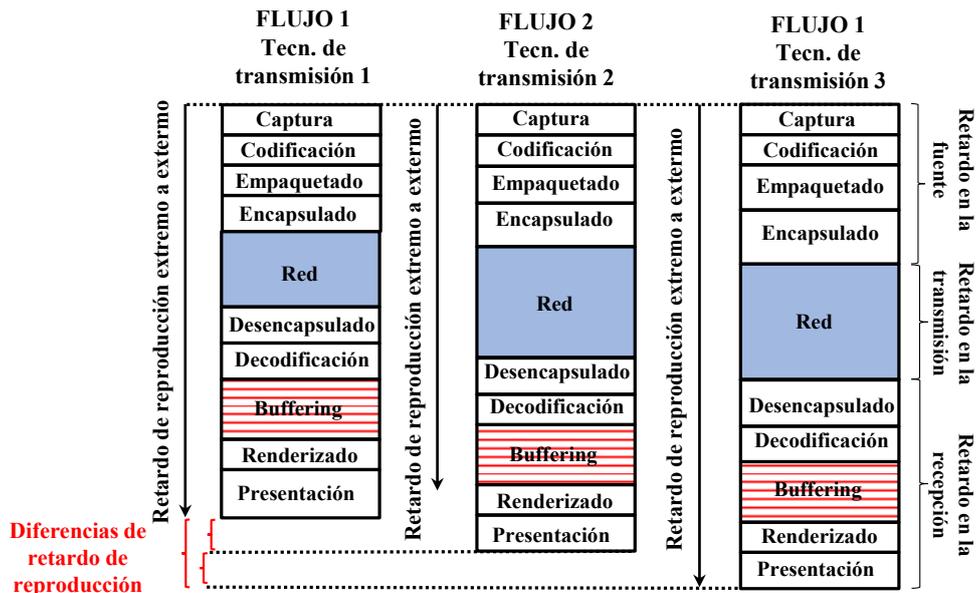


Fig. 2 Variabilidad de retardos extremo-a-extremo

Esta variabilidad de retardos se debe, principalmente, a los diferentes factores relacionados con la tecnología de transmisión, la red de distribución, o incluso al rendimiento y especificaciones de los dispositivos de generación y/o de consumo. La captura, codificación, empaquetamiento, la distribución por la red (influenciada por la carga de tráfico en la misma, los procesos de transcodificación -si fueran necesarios-, los de fragmentación y reensamblado de paquetes, las estrategias de encaminamiento dinámico o multicast, las políticas de colas implementadas en los routers, etc.), de procesado, desempaquetado, decodificación, renderizado y de presentación de los contenidos en el receptor, pueden alterar las relaciones temporales originales entre los diferentes flujos. Es decir, las relaciones temporales que existían entre las fuentes en el momento de su generación. Además, pueden entrar en juego otros problemas, como la pérdida de paquetes o la congestión de la red de transmisión, con lo que el retardo puede verse incrementado.

La existencia de dichos factores resulta en retardos extremo-a-extremo variables y diferentes, cuando la entrega de contenido se hace a través de diferentes tecnologías de transmisión a diferentes dispositivos. De la misma manera, el mismo contenido puede ser transmitido a través de diferentes tecnologías o redes (por ejemplo, en la Fig. 2, el flujo 1 es enviado a través de las tecnologías 1 y 3).

Tal variabilidad de retardos debe ser corregida o compensada en escenarios híbridos de consumo de contenidos multimedia mediante mecanismos de sincronización. Dichos mecanismos deben incluir especificaciones de diferentes técnicas, tales como la inserción e interpretación de líneas temporales (*timelines*), políticas de *buffering*, algoritmos de monitorización, técnicas de ajuste, así como de sincronización de relojes, tal y como se explica en la presente Tesis Doctoral.

2.4 NECESIDAD DE SINCRONIZACIÓN HÍBRIDA (**HYBRID SYNCHRONIZATION**)

Con el fin de alcanzar un nivel de sincronización adecuado en la presentación de contenidos multimedia recibidos por tecnologías híbridas (es decir, que las asincronías no sean perceptibles por el consumidor), es necesaria la manipulación y el ajuste de los procesos de reproducción de los contenidos involucrados. En el contexto de esta Tesis Doctoral, se denominan soluciones de **sincronización híbrida** al conjunto de procesos involucrados en conseguir la sincronización de los procesos de reproducción de los contenidos híbridos, bien en un único dispositivo o en varios. La sincronización híbrida es un caso particular de la sincronización conocida como sincronización inter-media o inter-flujo (es decir, cuando los diferentes contenidos a sincronizar, audio, video, etc., se reciben en flujos separados).

Por un lado, si en la cadena de distribución externo-a-extremo, se considera el lado de los proveedores de contenidos, los M flujos multimedia involucrados pueden ser generados y/o transmitidos de varias formas: 1) por un único proveedor; 2) por varios proveedores que compartan cierta relación o información entre ellos, como, por ejemplo, una misma referencia de reloj común a ser utilizada para generar las marcas de tiempo incluidas (*timestamps*) en los contenidos a transmitir; o 3) por proveedores totalmente independientes, incluyendo la situación en la que algún tercero realice una transcodificación o transmultiplexación del contenido multimedia original en algún punto de la cadena de distribución de contenidos externo a extremo.

Por otro lado, si se considera en dicha cadena el extremo del consumidor final, si hay múltiples dispositivos de consumo involucrados ($D > 1$) es necesaria la sincronización de la reproducción de contenidos en todos ellos. En el caso de que dichos dispositivos estén cercanos unos a otros (por ejemplo, en un escenario multipantalla en una red de área local, como una TV y una tablet o un smartphone) sería necesaria la sincronización interdispositivo o IDES. En el caso en el que los

dispositivos de consumo involucrados no estén próximos unos a otros (es decir, dispersos geográficamente), se empleará la sincronización interdestinatario o IDMS. Independientemente del caso en particular, al existir una amplia heterogeneidad en el ecosistema multimedia, garantizar la sincronización híbrida entre dispositivos es un reto que está cobrando especial relevancia en la actualidad. Primero, los flujos involucrados pueden contener cualquier tipo de contenido multimedia (audio, vídeo, texto, metadatos multisensoriales, etc.) con diferentes formatos y/o requisitos (por ejemplo, relativos al procesado o al ancho de banda). Segundo, las redes híbridas involucradas pueden diferir en términos de tipo de transmisión (unicast, multicast o broadcast), retardos intrínsecos a la tecnología utilizada, provisión de líneas temporales, disponibilidad de canales de comunicación bidireccionales, etc. Tercero, los dispositivos de consumo pueden diferir en términos de capacidad de procesamiento multimedia, recursos disponibles, especificaciones técnicas, etc. Concretamente, las magnitudes de retardos y su variabilidad son problemas muy relevantes en el ámbito de la transmisión y entrega de contenido multimedia híbrido, especialmente en la distribución broadband, debido al jitter (variabilidad del retardo) introducido por la propia red y por los dispositivos finales involucrados. De hecho, estos retardos se originan en los diferentes procesos existentes a lo largo de la cadena de transmisión extremo-a-extremo del contenido multimedia (Fig. 2), y pueden variar en diferentes instantes de tiempo, entre diferentes fuentes y dispositivos de consumo [17] [18] [19] [20].

Existen trabajos previos que han estudiado y analizado la magnitud de dichos retardos. Por ejemplo, en [17] se observó que en escenarios de IPTV las diferencias de retardos extremo-a-extremo pueden llegar a ser de hasta 6s y que las tecnologías broadband basadas en el uso de técnicas HAS podían introducir más de 60s de retardo respecto a las tecnologías broadcast. En dicho trabajo, también se detectó que existían diferencias significativas de retardo entre receptores, incluso cuando dichos receptores recibían los contenidos mediante la misma tecnología de distribución, el mismo servicio del proveedor (tipo de suscripción/calidad) y empleaban los mismos dispositivos de consumo finales. De todo ello, se puede deducir la necesidad de adoptar técnicas adaptativas y precisas de sincronización híbridas para compensar (y ajustar) los diferentes retardos asociados a este tipo de escenarios.

En las siguientes subsecciones se exponen los aspectos y componentes tecnológicos más relevantes que permiten que las técnicas de sincronización sean adaptativas y precisas.

2.5 SINCRONIZACIÓN DE RELOJES

Un requisito necesario para poder alcanzar un estado de sincronización satisfactorio en los entornos híbridos en los que se centra la presente Tesis Doctoral consiste en disponer de una noción del tiempo idéntica o, al menos, coherente entre las entidades involucradas. Por un lado, es necesario que los diferentes proveedores de contenido involucrados inserten líneas temporales comunes (o relacionadas de alguna manera) con el objetivo de poder alinear en el tiempo el consumo de los flujos transmitidos por cada uno de ellos. Por otro lado, también es necesario que los dispositivos de consumo involucrados puedan contar con información temporal consistente y robusta, para poder ajustar y sincronizar de manera precisa los procesos de reproducción involucrados.

Existen diferentes alternativas que pueden ser adoptadas para sincronizar relojes. La solución más utilizada en redes IP es el protocolo *Network Time Protocol* (en adelante, *NTP*), definido en la *RFC 5905* [21]. Dicho protocolo opera adoptando un modelo jerárquico cliente/servidor y se basa en la disponibilidad de servidores NTP de diferentes niveles (*stratums*) de precisión temporal, accesibles a través de redes IP, con los que sincronizarse. NTP es un protocolo soportado por un gran número de plataformas y equipamiento de red y, además, existen muchos servidores NTP

disponibles en Internet. Sin embargo, NTP asume retardos simétricos en los enlaces de la red, lo cual no es (demasiado) realista en las redes actuales. No obstante, mediante el uso de este protocolo, se pueden obtener niveles de sincronización de relojes de una precisión de decenas de milisegundos. Como alternativa, existe el protocolo *Precision Time Protocol* (en adelante, PTP), que está especificado en IEEE 1588-2008 [22] y que puede proporcionar niveles de sincronización de relojes más precisos que NTP. Sin embargo, para poder emplear PTP es necesario que los dispositivos involucrados sean compatibles con este protocolo. En este contexto, cabe destacar que en el *World Wide Web Consortium* (W3C), existe un grupo de trabajo denominado *Multi-Device Timing Community Group*, encargado de proporcionar un modelo de programación que permita compartir una línea de tiempo común en escenarios multidispositivo, basándose para ello en tecnologías Web y empleando un servidor externo. Esta solución también puede proporcionar niveles de sincronización precisas del orden de decenas de milisegundo, tal y como se puede comprobar en [23]. Por otro lado, también es posible implementar mecanismos de sincronización de relojes ad-hoc, como el implementado en el *framework* GStreamer [24] o el utilizado en [8]. De hecho, este tipo de mecanismos puede ser el que mejor encaje en escenarios multipantalla, puesto que los retardos entre los dispositivos involucrados son normalmente muy bajos. En estos escenarios, contar con un servidor externo remoto encargado de proporcionar una línea temporal común a todos los dispositivos puede no ser la mejor opción, ya que, cuanto más pequeños son los valores de latencia, más pequeños son los errores en los cálculos de retardo y, por tanto, la precisión que se puede alcanzar en la sincronización de relojes es mayor.

Adicionalmente, el consorcio Wi-Fi Alliance está trabajando en un mecanismo para proporcionar una sincronización precisa con valores de asincronías por debajo del microsegundo, llamado Wi-Fi Certified Sync⁵. Sin embargo, las últimas noticias son del año 2017 y no ha habido ninguna actualización posterior al respecto. A pesar de esto, dicho mecanismo podría suponer una solución potencial en escenarios donde la sincronización IDEs sea necesaria.

2.6 INSERCIÓN DE LÍNEAS TEMPORALES EN EL FLUJO BROADCAST

El flujo de transporte de MPEG2 (en adelante, MPEG2-TS) constituye la base del ecosistema actual de transmisión de contenido de TV digital broadcast por redes basadas en la tecnología DVB. MPEG2-TS especifica la forma en la que se contiene la información de los diferentes flujos de audio, vídeo y de datos (denominados flujos elementales, *Elementary Streams* o ES) de uno o varios canales de TV (denominados programas, *programs* según la terminología de MPEG2-TS) en paquetes. Al conjunto de los paquetes de cada flujo elemental se le denomina *Packetized Elementary Stream* o PES. Los PES se multiplexan en un único flujo de transporte continuo (*Transport Stream* o TS) formado por paquetes de tamaño fijo (denominados *transport packets*), de 188B. Para asegurar una sincronización correcta entre los contenidos transmitidos, en las cabeceras de los paquetes se insertan referencias de reloj y marcas temporales. Concretamente, la referencia temporal se proporciona a través del campo *Program Clock Reference* (en adelante, PCR) en la cabecera de los *transport packets*, mientras que las marcas temporales se almacenan en la cabecera de los paquetes de cada flujo en los campos denominados *Decoding, Presentation* y *Composition Timestamps* (en adelante, DTS, PTS y CTS, respectivamente), dependiendo del tipo de codificación adoptada (DTS/PTS si se utilizó MPEG2 o bien DTS/CTS si se utiliza MPEG4).

Este tipo de mecanismos proporciona referencias temporales relativas intrínsecas para sincronizar la información contenida dentro del mismo flujo de transporte MPEG2-TS. Sin embargo, su valor no tiene utilidad más allá de la sincronización de los contenidos de los

⁵ <https://www.wi-fi.org/discover-wi-fi/wi-fi-timesync>

programas incluidos en dicho flujo (sincronización intra- e inter-flujo). Si se quiere sincronizar la presentación de contenidos de programas incluidos en dos o más MPEG2-TS diferentes, los proveedores de dichos contenidos deben, necesariamente, emplear una misma referencia de reloj común (es decir, un reloj global) para insertar este tipo de marcas temporales relativas basadas en dicha referencia común o, al menos, proporcionar información que permita mapear las marcas de tiempo insertadas en los paquetes a dicha referencia de tiempo global común.

Otra limitación en cuanto al empleo de los campos de las marcas temporales incluidos en los flujos MPEG2-TS consiste en que las relaciones temporales contenidas en dichos campos pueden verse editadas y modificadas por equipos intermedios (multiplexores, transcodificadores, etc.) dentro de la cadena de distribución extremo-a-extremo, sin que el proveedor de dichos contenidos tenga conocimiento de dichas modificaciones. Esto puede afectar, de manera directa, dificultando, e incluso impidiendo, los procesos encargados de conseguir una sincronización entre contenidos que se basen en dichas marcas de tiempo y una referencia temporal común.

Para superar las limitaciones mencionadas, se han ido añadiendo extensiones a la especificación de MPEG2-TS, con el fin de proporcionar información temporal extrínseca y absoluta. En un primer momento, el Instituto Europeo de Estándares de Telecomunicaciones (*European Telecommunications Standards Institute*, o ETSI), propuso la especificación ETSI TS 102 823 [25], donde se define un mecanismo que permite insertar líneas temporales en el flujo de transporte MPEG2-TS. Dicho mecanismo consiste en la inserción de un descriptor de línea temporal (en adelante, *broadcast timeline descriptor*) y otro descriptor llamado *TV-Anytime* (en adelante, *TV-A*), dentro de la carga útil de los paquetes de un nuevo flujo de datos generado con información temporal, llamado *Timeline ES*, que será multiplexado junto con los otros ES en el MPEG2-TS (Fig. 3). Este nuevo flujo debe ser indicado en la tabla *Program Map Table* (PMT) como flujo de datos privados (*stream_type=0x06*, *stream_id=0xBD*). Para generar dicho flujo, en cada trama I del flujo de vídeo tomado como referencia, se genera un paquete del *Timeline ES* incluyendo información temporal, concretamente el valor del campo PTS de dicha trama I junto con el *broadcast timeline descriptor*, que contiene una marca de tiempo según una referencia de tiempo global.

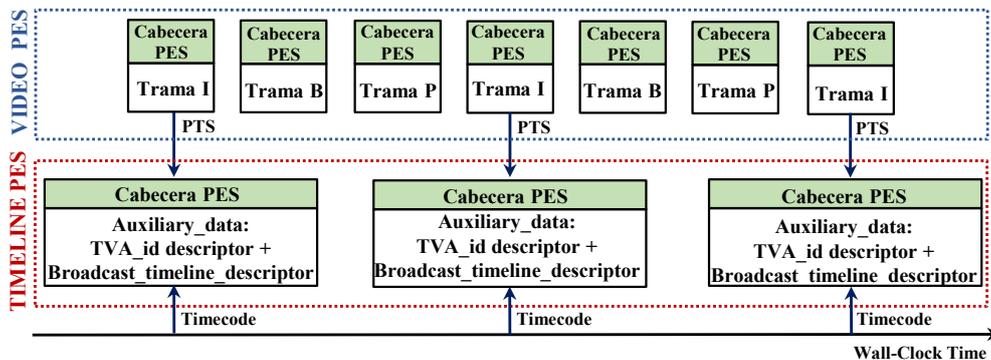


Fig. 3 Mecanismo de inserción de marcas temporales especificado por la especificación ETSI TS 102 823

Como tanto los paquetes de los flujos (PES) de Vídeo y Timeline contienen el mismo valor de PTS, se puede alcanzar una sincronización híbrida precisa, aunque el PCR sufra discontinuidades. Esto se consigue debido a que las líneas temporales extrínsecas que se han insertado por parte del proveedor de contenidos en el flujo de transporte no serán alteradas por ningún elemento en la cadena de distribución, al ser parte del contenido MPEG2-TS. Por tanto, este mecanismo permite incluso sincronizar contenidos en casos en los que haya diferentes proveedores de contenido involucrados. Además, este tipo de líneas temporales son agnósticas respecto al contenido y a la tecnología de transmisión, por lo que también pueden adoptarse para sincronizar

flujos de vídeo bajo demanda o de datos o eventos particulares, que normalmente no incluyen ningún tipo de información para poder establecer mecanismos de sincronización.

Más recientemente, se ha propuesto otra solución más flexible y eficiente⁶ en términos de ancho de banda (en adelante *BW*, del inglés *bandwidth*) con respecto a la anterior solución descrita. Ha sido propuesta por MPEG y DVB como una corrección al ISO/IEC 13818-1, con el nombre de *Timing and External Media Information* (en adelante, *TEMI*) [26]. En particular, el mecanismo TEMI proporciona las siguientes características:

- Inserción de líneas temporales extrínsecas, absolutas y estables en el flujo de transporte MPEG2-TS.
- Inserción de URL, indicando la ubicación de contenido relacionado y disponible a través de redes de banda ancha, en el flujo de transporte MPEG2-TS.
- Notificación de cuándo el contenido adicional va a estar disponible, pudiendo transmitir señales de cuenta atrás para una línea temporal determinada.

Los dos primeros tipos de metadatos se conocen como descriptores *TEMI timeline* y *TEMI location*, respectivamente. Estos descriptores pueden insertarse (con una tasa de inserción configurable) dentro del flujo MPEG2-TS de dos formas distintas (Fig. 4): i) en el campo de adaptación (*Adaptation Field o AF*) incluido en la cabecera de los paquetes de transporte MPEG2-TS (*transport packets*) del flujo principal (por ejemplo, el de vídeo) (Fig. 5); o ii) en los paquetes de un flujo de datos (PES) dedicado, que deberá ser indicado en la PMT como de datos privados (*stream_type=0x26, stream_id=0xBD*) y multiplexado en el flujo MPEG2-TS. La primera opción es la más eficiente en términos de uso de ancho de banda, motivo por el cual se adopta en esta Tesis Doctoral

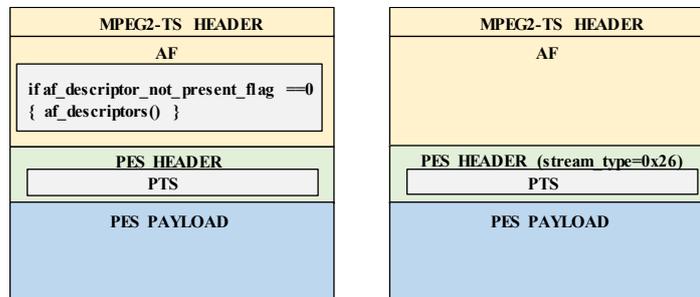


Fig. 4 Dos opciones de inserción de los descriptores Timing and External Media Information

⁶ De acuerdo con [34], para un flujo de vídeo a 50Hz, la inserción de una línea temporal (timeline) utilizando el mecanismo propuesto por la ETSI [25] introduce una sobrecarga de 75Kbps. Por el contrario, adoptando el mecanismo TEMI [26] se reduce a 4Kbps.

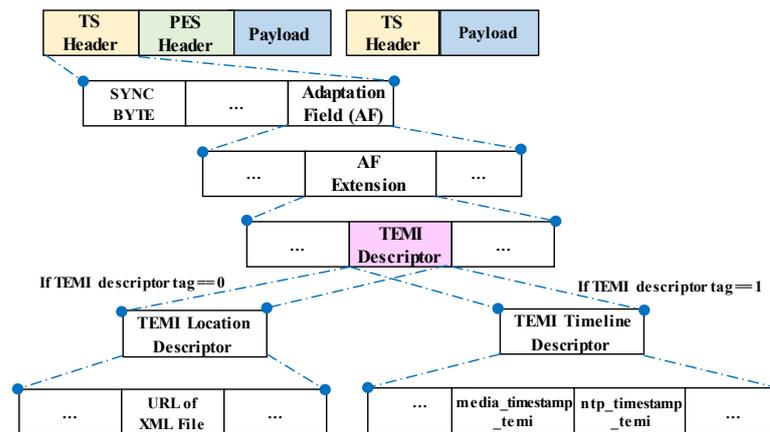


Fig. 5 Inserción de los descriptores TEMI en el campo de adaptación de los paquetes de transporte

2.7 ARRANQUE Y ADQUISICIÓN DE CONTENIDOS HÍBRIDOS

Para conseguir una sincronización híbrida en uno o varios dispositivos (o destinatarios), es necesario proporcionar una serie de mecanismos adicionales que deben ser ejecutados durante el proceso de arranque de las aplicaciones de reproducción, para así adquirir los contenidos adicionales y complementarios al contenido principal. Por un lado, durante este proceso de arranque, es necesaria una etapa durante la cual los dispositivos involucrados adquieren y lanzan las aplicaciones específicas que permiten identificar e iniciar la recepción de los diferentes flujos híbridos relacionados disponibles. Por otro lado, el ajuste de la reproducción de contenidos híbridos adicionales relacionados se refiere al proceso mediante el cual las dependencias temporales/espaciales y/o semánticas de los diferentes contenidos relacionados se ajustan para poder proporcionar una experiencia de consumo sincronizada satisfactoria.

En los servicios de difusión tradicionales (broadcast), la información para el proceso de arranque se puede obtener bien mediante estrategias *off-band* o bien estrategias *in-band*. Un ejemplo del primer tipo es el protocolo *Session Description Protocol* (en adelante, SDP) [27], mientras que un ejemplo del segundo tipo de estrategia, es el empleo de la tabla PMT (*Program Mapping Table*) cuando se utilizan flujos de transporte MPEG2-TS. En la distribución de contenidos híbridos, esta última opción no es viable, puesto que se necesita una asociación entre los flujos multimedia involucrados provenientes de diferentes fuentes. Una alternativa consiste en señalar la existencia de contenido multimedia adicional, disponible vía broadband, a través del protocolo SDP, compartiendo una Guía Electrónica de Programa (en adelante, EPG, del inglés *Electronic Program Guide*) o cualquier otro mecanismo propietario.

Para proporcionar la información necesaria para el proceso de arranque, se pueden adoptar dos estrategias diferentes, basadas bien en servidor (*server-based*) o bien en el flujo multimedia (*stream-based*). En la estrategia basada en servidor, las aplicaciones del consumo de contenidos híbridos consultan a un servidor central (por ejemplo, alojado por el propio difusor de contenidos o *broadcaster*) acerca de la ubicación e información necesaria (metadatos) de los contenidos multimedia adicionales, disponibles a través de redes broadband. El servidor les responde con la información correspondiente, a través de alguno de los métodos comentados (por ejemplo, SDP o EPG). En la estrategia basada en flujo multimedia, dicha información puede insertarse en el propio flujo MPEG2-TS transmitido por la red broadcast (por ejemplo, a través de descriptores que contengan las URL de ficheros conteniendo la información relativa a los contenidos

multimedia adicionales disponibles a través de redes broadband, como, por ejemplo, su ubicación o formato). En esta Tesis Doctoral, se decidió utilizar esta última estrategia puesto que permite obtener de manera instantánea cualquier actualización al respecto de cualquier cambio asociado a la información insertada por parte del proveedor o encargado de la generación de contenidos, sin necesidad de llevar a cabo una monitorización periódica, a diferencia de la estrategia basada en servidor.

2.8 DESCUBRIMIENTO Y COMUNICACIÓN ENTRE DISPOSITIVOS EN UN ENTORNO LOCAL

Otros aspectos esenciales para la sincronización de contenidos son el descubrimiento y enlazado de sus procesos de reproducción. Estos procesos de reproducción pueden estar ejecutándose bien en el mismo terminal híbrido (donde también se ejecuta la aplicación Main Screen, en adelante denominada MS app), bien en diferentes dispositivos secundarios (donde se ejecuta la aplicación Companion Screen, en adelante denominada CS app). El objetivo del proceso de descubrimiento y enlazado de ambos procesos (MS y CS app) es notificar la existencia de contenidos relacionados, para ser reproducidos de manera sincronizada. Por lo tanto, el primer proceso a realizar deberá ser el de descubrimiento, bien de dispositivos o bien de procesos (en escenarios multidispositivo, como, por ejemplo, en aplicaciones multipantalla, o en escenarios con un único terminal híbrido, como, por ejemplo, en una Smart TV, respectivamente), ver Fig. 6.

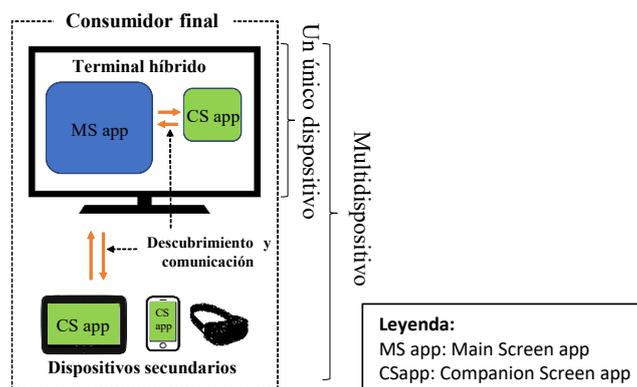


Fig. 6 Necesidad de descubrimiento y comunicación en un escenario mono y multidispositivo

En particular, para escenarios multidispositivo en un entorno local, en los cuales se deberá implementar una solución de sincronización IDEs, respectivamente, es necesario un mecanismo adicional para la vinculación entre los dispositivos involucrados, así como el establecimiento de un canal de comunicación entre los mismos. Dicho mecanismo permitirá a los dispositivos conocer la existencia de otros dispositivos y, de esta forma, poder crear una sesión sincronizada compartida, estableciendo un canal de comunicación entre todos ellos para intercambiar la información necesaria para poder interactuar entre ellos y permitir alcanzar un nivel de sincronización aceptable.

Respecto a dicho proceso de descubrimiento de dispositivos en escenarios multidispositivo, se pueden emplear soluciones que implican interacción por parte de los usuarios, como, por ejemplo, introducir un código PIN, escanear un código QR o un código de barras, etc. Sin embargo, es deseable que el proceso de descubrimiento y vinculación de dispositivos sea más automático y transparente para el usuario, sin necesidad de que este intervenga de forma explícita. En el caso de escenarios multidispositivo locales, tecnologías como *Universal Plug and Play* (UPnP) [28] o *Digital Living Network Alliance* (DLNA) [29] podrían ser candidatas para dicha

finalidad. Sin embargo, no están adoptadas ni amplia ni uniformemente (lo que podría comportar problemas de interoperabilidad) y su implementación requeriría de un esfuerzo significativo. Una alternativa más apropiada podría consistir en utilizar una combinación de DNS Multicast (multicast DNS o mDNS) junto con el Servicio de Descubrimiento DNS (*DNS-Service Discovery* o DNS-SD), propuesto en [8] y también conocido como *Zeroconf*.

2.8.1 DISCOVERY AND LAUNCH PROTOCOL

El protocolo Discovery and Launch (DIAL) [30] se ha propuesto como una herramienta sencilla para escenarios multidispositivo, que permite a los dispositivos (o procesos) secundarios (Companion Screens o CS) descubrir dispositivos (o procesos) principales (Main Screens o MS) y lanzar aplicaciones en los mismos. En concreto, este mecanismo ha sido desarrollado por Netflix y Youtube, y ha sido adoptado como el mecanismo de descubrimiento y asociación de dispositivos por otras soluciones, como las propuestas en el estándar HbbTV [5] (explicado en profundidad, más adelante). Por este motivo, también es el mecanismo de descubrimiento adoptado en la solución propuesta en esta Tesis Doctoral. En particular, DIAL ofrece un mecanismo para la asociación de dispositivos totalmente transparente para el usuario y, además, proporciona un canal de comunicación entre los dispositivos (o procesos) MS y CS disponibles, permitiendo una integración inter-dispositivo (o inter-proceso) satisfactoria. Por un lado, este mecanismo se basa en el protocolo *Simple Service Discovery Protocol* (SSDP) para el proceso de descubrimiento de dispositivos locales. Dicho mecanismo está integrado en el protocolo UPnP [31]. Por otro lado, para el canal de comunicación establecido entre los dispositivos, se adoptan tecnologías basadas en protocolos HTTP, como el protocolo Websocket (WS) [32]. El mecanismo DIAL se basa en un esquema cliente/servidor, y está compuesto de dos partes:

- Servicio de Descubrimiento DIAL (*DIAL Service Discovery*): permite a un *cliente DIAL* descubrir a otros dispositivos que se encuentren en la misma red de área local, y obtener acceso al servicio REST de DIAL en dichos dispositivos.
- Servicio REST de DIAL (*DIAL REST Service*): permite que un cliente DIAL consulte, lance y, opcionalmente, termine aplicaciones activas en un servidor DIAL.

2.8.2 COMPANION SCREENS AND STREAMS

En escenarios con uno o múltiples dispositivos (o procesos) de reproducción, se asume que el tipo de comunicación que más fiabilidad y rapidez puede ofrecer para alcanzar un estado de sincronización satisfactorio consiste en una comunicación directa entre los mismos. El grupo Digital Video Broadcasting (DVB), que está compuesto, entre otros, por fabricantes de TV, operadores de radiodifusión y/o de plataformas de TV, participó en la especificación de casos de uso y de los requisitos comerciales necesarios para poder ofrecer experiencias de TV con segundas pantallas, así como en las actividades de estandarización asociadas (a través del grupo DVB CM-COS⁷, actualmente inactivo). Fruto de dichas actividades, se propuso la especificación *DVB Companion Screens and Streams* (DVB-CSS) [33]. En particular, dicho documento define un marco de trabajo estandarizado que permite que dispositivos (o procesos) CS, junto a un dispositivo (o proceso) MS, implementen una experiencia de consumo de contenido multipantalla de manera sincronizada, todo ello a través de la red de área local del propio hogar. DVB-CSS incluye una serie de mecanismos que permiten que el MS pueda notificar a los CS involucrados qué contenido está reproduciendo el MS en un determinado instante y la referencia de tiempo global actual para dicho instante, permitiendo la sincronización entre los procesos de reproducción que se ejecutan en el MS y en el CS, respectivamente. DVB-CSS considera el uso del mecanismo TEMI [26] descrito, en combinación con los protocolos Websocket [32] y UPnP (y, por

⁷ <https://free.dvb.com/index.php/groups/CM-COS>

tanto, SSDP) [28]. En particular, la especificación DVB-CSS incluye diferentes interfaces, módulos y protocolos para conseguir diferentes propósitos, como son los siguientes [34]:

- Descubrimiento y asociación de dispositivos.
- Identificación de contenidos e intercambio de información de estado.
- Sincronización de relojes (*WallClock Sync*). El protocolo *Wallclock Sync* (CSS-WC) proporciona los medios necesarios para utilizar un algoritmo de sincronización de relojes para establecer una aproximación *best-effort* del *Wall Clock* del MS en las aplicaciones CS.
- Sincronización de la línea temporal (*Timeline Synchronization* o TS). Permite el intercambio de información temporal (timestamps) entre los dispositivos (MS y CSs) involucrados. El mecanismo de sincronización de TS (CSS-TS) es utilizado por el MS para informar sobre determinados instantes de tiempo en su línea temporal (*timeline*) al resto de CS involucrados. Dichos instantes pueden estar señalizados en el flujo broadcast o pueden ser deducidos del formato del contenedor del flujo (*streaming container format*).
- Disparo de eventos (*event triggering*). Detección y ejecución de eventos como, por ejemplo, los conocidos como *Do it now events* del flujo broadcast. Dichos eventos son utilizados en plataformas de aplicaciones basadas en DVB, como MHP (*Multimedia Home Platform*) o las plataformas HbbTV, y no tienen asociado ningún instante de ejecución, sino que deben ser ejecutados tan pronto como sean recibidos.

En el desarrollo de esta Tesis Doctoral, se han adoptado los módulos de descubrimiento y asociación de dispositivos, de identificación de contenidos e intercambio de información de estado, y de sincronización de la línea temporal, con el fin de que exista un canal de comunicación entre dispositivos MS y CS y puedan transmitirse la información necesaria para proporcionar un escenario IDES.

2.9 ESTÁNDARES QUE CONTEMPLAN LA RECEPCIÓN Y EL CONSUMO DE CONTENIDOS HÍBRIDOS

En esta sección se exponen los estándares más relevantes capaces de recibir, gestionar y presentar contenido híbrido.

2.9.1 HYBRID BROADCAST BROADBAND TV

El protocolo HbbTV [5] es una iniciativa global con el objetivo de armonizar la entrega de contenidos broadcast y contenidos broadband en servicios de entretenimiento a los consumidores, a través de dispositivos conectados, como TV conectadas, *Set-Top Boxes* (STBs) o dispositivos secundarios, como, por ejemplo, móviles o tablets. En sus últimas versiones (la actual es la v2.0.3), este estándar especifica nuevas funcionalidades para alcanzar una sincronización híbrida en un único dispositivo (es decir, un terminal híbrido) y/o entre múltiples dispositivos (es decir, escenarios IDES). Las características más relevantes que ofrece este estándar, en el contexto de esta Tesis Doctoral, son las siguientes:

- Hace uso de servicios DVB-T/S broadcast convencionales (MPEG2-TS). A pesar de que se debe añadir una serie de extensiones en los flujos DVB, el estándar HbbTV asegura la retrocompatibilidad con aquellos dispositivos que no implementen el estándar.
- Adopta tecnologías web actuales, tales como HTML5, JavaScript, CSS y Websockets.

- Adopta MPEG DASH [4] como la tecnología de streaming adaptativo de contenidos multimedia disponibles a través de redes broadband.
- Adopta el protocolo DIAL [30] para el descubrimiento (a través del protocolo SSDP) entre dispositivos (o procesos) CS y MS en escenarios multipantalla. Además, contempla el uso del DIAL para el lanzamiento y ejecución de aplicaciones desde/hacia segundas pantallas.
- Servicio de comunicación denominado *App-to-App*, a través de un servidor de Websockets (WSS o WebSocket Server), con el que establecer un canal de comunicación directa entre el dispositivo (o proceso) MS y los dispositivos (o procesos) CS involucrados.
- Se propone el uso de la tabla *Application Information Table* (AIT) incluida en el flujo DVB, para señalar la existencia de aplicaciones HbbTV disponibles. En dicha tabla se almacena la URL donde está alojada la aplicación que debe ejecutarse al iniciar la reproducción de un determinado programa del MPEG2-TS (es decir, un canal de TV) recibido.
- Sincronización (híbrida) entre flujos multimedia relacionados o entre flujos multimedia y aplicaciones en un solo terminal híbrido. Para estos casos, HbbTV adopta la solución TEMI descrita en este capítulo (apartado 2.6) para la inserción de marcas temporales y la información requerida.
- Sincronización (híbrida) entre flujos multimedia relacionados o entre flujos multimedia y aplicaciones en escenarios multidispositivo. Para estos casos, HbbTV adopta la especificación DVB-CSS.

En esta Tesis Doctoral, se proponen soluciones de sincronización híbrida compatibles con este estándar, por lo que, como se ha indicado anteriormente, se han adoptado algunos de los protocolos y especificaciones anteriormente definidas (por ejemplo, DIAL, WebSocket, DVB-CSS). De esta forma dichas soluciones pueden ser implementadas y desplegadas de forma satisfactoria, con tecnologías actualmente en servicio.

2.9.2 MULTIMEDIA HOME PLATFORM

El estándar Multimedia Home Platform (MHP) fue diseñado por el consorcio DVB y también ha sido estandarizado por la ETSI. MHP define una plataforma común para aplicaciones interactivas de la Televisión Digital, independiente del proveedor de contenidos y del receptor empleado. Está basado en Java y, a pesar de ser una tecnología con más años, ha quedado desplazada en favor de HbbTV. Esto es debido a diversos factores, siendo uno de ellos la flexibilidad de HbbTV (al emplear HTML5), ya que MHP necesita una máquina virtual Java para poder ejecutar las aplicaciones.

2.9.3 MPEG MEDIA TRANSPORT

MMT [35] es otro estándar que forma parte del ISO/IEC 23008 [36]. Se propone como el sustituto de transporte multimedia MPEG2-TS, con el objetivo de ofrecer soporte a nuevos servicios como la UHD TV (TV de ultra-alta resolución), segundas pantallas, etc., siendo compatible con las tecnologías web (HTML5). Este estándar permite la multiplexación de contenidos y su distribución por redes híbridas.

2.9.4 GINGA

Ginga es un middleware (es decir, un software situado en una capa intermedia entre el sistema operativo y las aplicaciones que se ejecutan en él) que está implantado en algunas regiones del mundo en las que no se ha adoptado el estándar HbbTV (como por ejemplo, Brasil o Japón). Se trata de una especificación nipo-brasileña, que permite el desarrollo de aplicaciones interactivas para TDT, independientemente del receptor y de los terminales de acceso. Su arquitectura consta

de tres módulos: Ginga-NCL, Ginga-J y Ginga-CC. El primero de los módulos provee una infraestructura de presentación para aplicaciones interactivas de tipo declarativas empleando el lenguaje NCL (*Nested Content Language*), basado en XML; el segundo provee aplicaciones basadas en Java; y, finalmente, el tercero es el Núcleo Común (*Common Core*) de los otros dos.

2.9.5 ADVANCED TELEVISION SYSTEMS COMMITTEE

ATSC 3.0 [37] es la última versión del estándar americano para servicios *broadcast*, que también soporta el envío coordinado a través de redes *broadband*, soportando DASH. En particular, este estándar también contempla el consumo de contenido multimedia a través de dispositivos heterogéneos, así como el uso de las redes *broadband* para la transmisión de contenido bajo demanda y una mayor interactividad, a través de tecnologías web (HTML5).

2.10 COMPONENTES DE LAS SOLUCIONES DE SINCRONIZACIÓN ENTRE FLUJOS MULTIMEDIA

En esta sección se presentan diferentes esquemas, estrategias y mecanismos que permiten el consumo de contenidos híbridos, así como soluciones para alcanzar un estado de sincronización satisfactorio, tanto en escenarios donde se requiere una sincronización multidispositivo tanto a nivel local (IDES), como a nivel remoto, en escenarios multidestinario (IDMS).

2.10.1 ESQUEMAS DE CONTROL DE LA SINCRONIZACIÓN

Existen diferentes esquemas de control de la sincronización de varios procesos de reproducción. Dichos esquemas se pueden dividir en dos tipos: centralizados y distribuidos. Por un lado, dentro de los esquemas de control centralizados se encuentran los esquemas *Master/Slave* (M/S) y el *Synchronization Maestro Scheme* (SMS). En el esquema M/S, los receptores reproductores de contenido se diferencian al adoptar un rol de máster (un solo receptor) o de esclavo (el resto de los receptores). En este esquema, el receptor máster envía, al resto de receptores de la misma sesión, mensajes con la información necesaria para que ajusten sus procesos reproductores y así poder alcanzar un estado de sincronización satisfactorio. Al recibir dichos mensajes, los receptores esclavos ajustan sus procesos de reproducción en base a la información temporal recibida. En cambio, el esquema SMS se basa en la existencia de una entidad (ya sea el transmisor del contenido, uno de los receptores involucrados o una entidad independiente), denominada *maestro*, que es la encargada de obtener la información de los procesos de reproducción de cada uno de los receptores involucrados, para calcular un valor que será utilizado como referencia temporal de sincronización. Además, este esquema permite que la entidad maestra pueda enviar una referencia común a todos los receptores de la sesión, o distribuir información personalizada a cada uno de ellos para que ajusten su proceso de reproducción debidamente. Ambos esquemas (M/S y SMS) se comportan de una manera similar, sin embargo, en el esquema M/S solamente el receptor con rol de máster envía mensajes (comunicación unidireccional), mientras que, en el esquema SMS, es necesario que todos los receptores envíen información acerca de sus respectivos procesos de reproducción y, en sentido opuesto, la entidad maestra les envíe mensajes con información para que ajusten sus procesos de reproducción (comunicación bidireccional).

Por otro lado, dentro de los esquemas de control distribuidos se encuentra el *Distributed Control Scheme* o esquema DCS. En dicho esquema, todos los receptores reciben información acerca de los procesos de reproducción del resto de receptores. De esta manera, cada uno de los receptores puede, de una manera independiente, calcular una referencia de sincronización propia con la que comparar su estado de reproducción y realizar ajustes en sus procesos de reproducción para sincronizarlo con dicha referencia.

La Fig. 7 muestra los tres tipos de esquemas de control posibles que permiten alcanzar un estado de sincronización entre procesos de reproducción, involucrando contenido multimedia y un canal de comunicación (uni- o bidireccional) para alcanzar la sincronización. Se puede observar cómo en todos los esquemas de control se envía información de los procesos de reproducción por parte del máster (en el esquema M/S); a la entidad maestro (en el esquema SMS); o de todos (a todos) los receptores (en el esquema DCS). Únicamente en el caso del esquema SMS se envía un mensaje diferente, que incluye la referencia a la cual ajustar los procesos de reproducción.

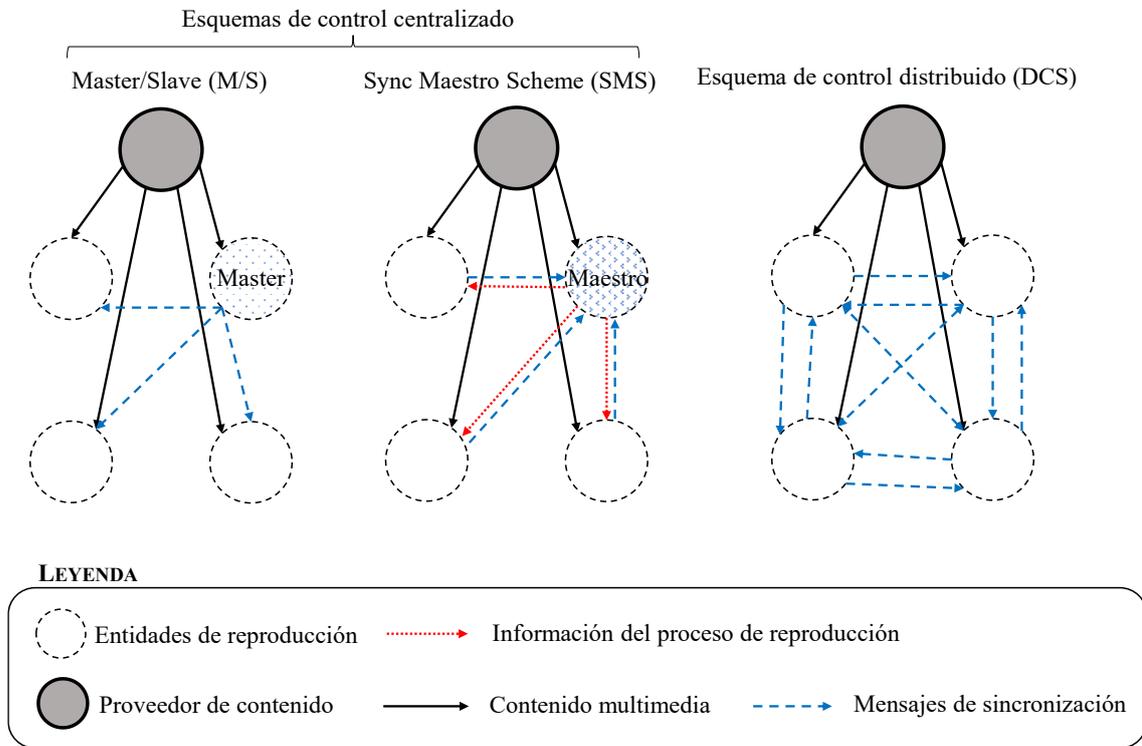


Fig. 7 Esquemas de control del proceso de sincronización entre procesos de reproducción

2.10.2 ESTRATEGIAS DE SELECCIÓN DE LA REFERENCIA DE SINCRONIZACIÓN

Para alcanzar un estado de sincronización satisfactorio entre los diferentes procesos de reproducción involucrados, pueden adoptarse diferentes políticas de sincronización [38], según cuál es o cómo se calcula la referencia a la cual se deben ajustar dichos procesos:

- Sincronización al más lento: en esta estrategia, la referencia pasa a ser el receptor cuyo proceso de reproducción se encuentra en el instante del contenido más atrasado, respecto al resto de receptores del mismo grupo o sesión.
- Sincronización al más rápido: en esta estrategia, la referencia pasa a ser el receptor cuyo proceso de reproducción se encuentra en el instante del contenido más adelantado, respecto al resto de receptores del mismo grupo o sesión.
- Sincronización a un punto intermedio: en esta estrategia, la referencia se calcula como el instante intermedio de todos los receptores del mismo grupo o sesión.
- Sincronización a la tasa nominal del servidor: en esta estrategia, el proveedor de contenidos o servidor actúa como un receptor "virtual", por lo que proporciona una referencia de ajuste del proceso de reproducción ideal al resto de receptores.

De estas cuatro estrategias de selección de la referencia de sincronización, solamente la cuarta puede ser seleccionada únicamente si el esquema de control adoptado es el SMS, para

el caso en el que la entidad maestra esté vinculada con el proveedor o la fuente emisora de contenidos. El esquema de control M/S puede adoptar cualquiera de las dos primeras estrategias (esto es, al más lento o al más rápido) y el esquema de control DCS puede adoptar cualquiera de las tres primeras estrategias (al más lento, al más rápido o a un punto intermedio).

2.10.3 TÉCNICAS DE AJUSTE EN EL PROCESO DE REPRODUCCIÓN DEL CONTENIDO

Existen diferentes técnicas para realizar ajustes en el proceso de reproducción de un contenido multimedia. Por un lado, se pueden realizar saltos y pausas (S&P) como una forma agresiva de realizar un ajuste. A pesar de ser un tipo de ajuste brusco y perceptible por el usuario, es una técnica comúnmente utilizada. Por otro lado, se pueden realizar otro tipo de ajustes más suaves, adoptando técnicas denominadas como *Adaptive Media Playout* (AMP) ([39], [40]). Este tipo de técnicas consisten en la modificación de la tasa de reproducción (incrementándola o reduciéndola de manera sutil) con el fin de alcanzar lo antes posible y mantener un nivel de sincronización dentro de niveles tolerables.

2.11 CONTENIDOS DE NUEVA GENERACIÓN

En esta sección se va a tratar de definir los dos tipos de contenidos que se pueden considerar de nueva generación, al tratarse de contenido que se extiende más allá del contenido tradicional, como son el contenido omnidireccional y el contenido multisensorial.

2.11.1 CONTENIDO OMNIDIRECCIONAL

En la actualidad, la generación de contenido multimedia ha evolucionado desde el vídeo 2D tradicional hacia un contenido más inmersivo, como el vídeo omnidireccional (que proporciona visión en los 360 grados de una escena) o incluso contenido volumétrico [41]. Estos nuevos tipos de contenido de vídeo requieren un mayor ancho de banda y tecnologías que soporten una mayor calidad de codificación.

El contenido omnidireccional, puede ser reproducido en pantallas planas como en otros dispositivos capaces de monitorizar la orientación del usuario, con el fin de que la visualización de dicho contenido se lleve a cabo de una forma más natural e intuitiva, denominados Head Mounted Displays o HMDs.

Durante el año 2020, la previsión del gasto global en dispositivos HMD fue de 10.6kM\$ [42]. Además, se espera que entre los años 2021-2027 el consumo de este tipo de dispositivos se incremente en un 36.3% [43]. De acuerdo con esta información, se espera que dispositivos como los HMD se vayan adoptando para el consumo de contenido omnidireccional disponible a través de Internet. De esta manera, también se puede proporcionar contenido de esta naturaleza para complementar el contenido multimedia recibido a través de redes broadcast (como es el caso del contenido de TV), permitiendo proporcionar experiencias de consumo más inmersivas y realistas. Como ejemplo, en España, el radiodifusor público RTVE ha habilitado en el pasado reciente contenido de realidad virtual (RV) e interactivo en series como el *Ministerio del Tiempo* o *El Caso. Crónica de sucesos*.

Además, una prueba de la creciente demanda de contenido omnidireccional es la cantidad de esfuerzos para codificar y transmitir este tipo de contenido de manera eficiente. Puesto que la transmisión de este tipo de contenido omnidireccional es un reto clave. Los usuarios solamente observan un pequeño porcentaje de toda la escena en 360° completa, denominado campo de visión o Field of View (FoV), tal y como se muestra en la Fig. 8.

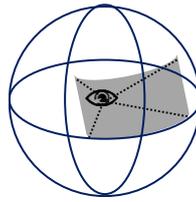


Fig. 8 Porción visualizada por el usuario en un contenido omnidireccional

Es por ello por lo que, si se transmitiera todo el contenido 360 disponible al usuario, supondría un uso ineficiente del ancho de banda disponible, por el hecho de que la proporción de pixels siendo visualizados por el usuario, comparado con el número de pixels totales transmitidos, es relativamente baja.

Otra evidencia de que el contenido omnidireccional (y su consumo a través de dispositivos HMD) está ganando protagonismo, son trabajos de investigación como [44], donde se buscan diversos mecanismos que permitan una mejor interacción entre usuarios que estén consumiendo simultáneamente el mismo contenido, pero a través de sendos dispositivos HMD, con el fin de evitar el aislamiento que estos dispositivos crean, como por ejemplo, chat por voz/audio, videochat o identificación del FoV del compañero.

Los contenidos multimedia omnidireccionales requieren, por un lado, de más ancho de banda y, por otro, de soporte de tecnologías que permitan codificar con mayor calidad. El estándar MPEG-I [45] especifica los tipos de contenido multimedia emergentes, con el objetivo principal de estandarizar contenido multimedia inmersivo. Además, se describen las características del contenido multimedia inmersivo, distinguiéndolo del contenido multimedia 2D tradicional. Algunas de estas características son: la generación de contenido 360 o estereoscópico (y el gran abanico de posibilidades y mecanismos existentes para crearlo), la adopción de nuevas técnicas de codificación para contenido de alta resolución (como, por ejemplo, *High Efficiency Video Coding* o HEVC [36]), la utilización de dispositivos específicos para su visualización o consumo (como, por ejemplo, los dispositivos HMD), etc.

Respecto a la capacidad de consumir contenido de vídeo 360, existen diferentes niveles de libertad (*Degrees of Freedom* o DoF). Como ejemplo, un contenido de vídeo omnidireccional con 3 grados de libertad (3DoF), permite al usuario consumir el contenido con un simple HMD y sentado en una silla. De hecho, con una simple rotación de la cabeza en cualquiera de los 3 ejes representados en la Fig. 9, se pueden consumir diferentes secciones del contenido. En la misma se aprecian los tres movimientos de rotación, conocidos como viraje, inclinación y balanceo (yaw, pitch y roll, respectivamente).

El contenido de vídeo omnidireccional con 3DoF es el más común, pero no es suficiente para proporcionar al consumidor una experiencia totalmente inmersiva. De hecho, se ha definido el concepto de 3DoF+ como evolución respecto al de 3DoF [46]. En particular, el contenido de vídeo omnidireccional 3DoF+ permite movimientos (limitados) a través de los ejes de viraje, inclinación y balanceo, por lo que el contenido debe proporcionar más información respecto a la profundidad y texturas del contenido. Finalmente, existe el contenido omnidireccional 6DoF (Fig. 9), en el que el consumidor se puede mover "libremente" por la escena virtual sin ningún tipo de limitación, es decir, no sólo mediante rotación en los 3 ejes indicados (3 grados de libertad) sino también mediante traslación a lo largo de los mismos (otros 3 grados de libertad).

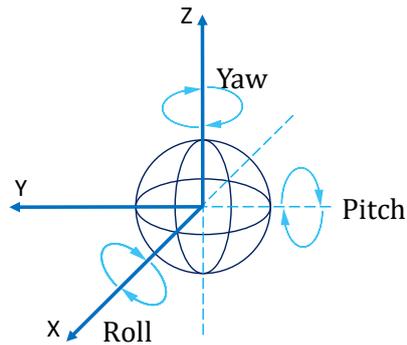


Fig. 9 Viraje (yaw), inclinación (pitch) y balanceo (roll), además de la traslación en los ejes X, Y, Z.

TIPOS DE PROYECCIÓN PARA CONTENIDO OMNIDIRECCIONAL

Cuando se genera contenido de vídeo omnidireccional a través de cámaras 360, la metodología a partir de la cual las imágenes capturadas se proyectan en un plano bidimensional no es evidente [47]. Durante el proceso de generación, pueden suceder ciertos errores como cierta distorsión en las imágenes 360. Como ejemplo para evitar este tipo de problemas, en [47] se propone la combinación de múltiples proyecciones locales con múltiples puntos de vista diferentes para crear proyecciones más realistas. Además, en [48] se presenta un estudio exhaustivo de los diferentes tipos de proyecciones existentes y de las diferentes metodologías para evaluar su calidad. En dicho trabajo, se clasifican las proyecciones en dos categorías: dependientes del punto de vista o independientes del punto de vista. Dentro de las proyecciones independientes del punto de vista, se encuentran la proyección equirrectangular (ERP) o la proyección de mapeado cúbico (*cube map* o CMP). En cambio, dentro de las proyecciones dependientes del punto de vista, donde se tiene en cuenta la región de interés del consumidor (es decir, el FoV), se encuentra la proyección piramidal, propuesta por Facebook en [49]. La Fig. 10 muestra un ejemplo de las tres proyecciones descritas.

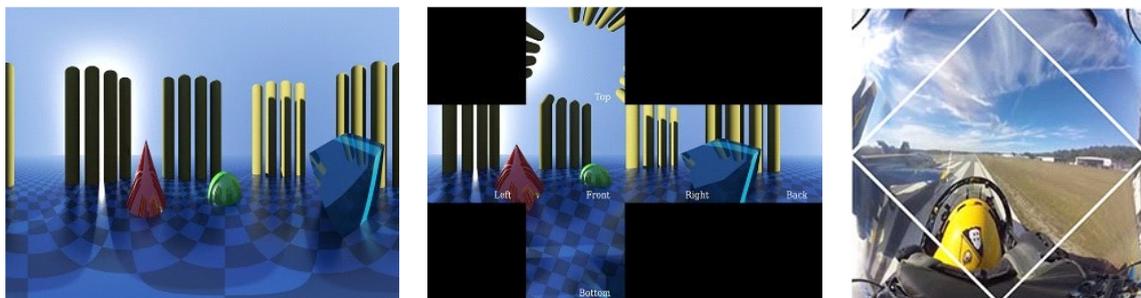


Fig. 10 Proyecciones más comunes en vídeo 360: ERP (izq.), CMP (centro) y piramidal (dch.)

TRANSMISIÓN DE CONTENIDO OMNIDIRECCIONAL

Tal y como se ha indicado anteriormente, la transmisión de contenido omnidireccional requiere un mayor uso de ancho de banda, en comparación con la transmisión de contenido de vídeo 2D tradicional. Teniendo en cuenta que el consumidor únicamente visualiza una pequeña región del contenido (el FoV), se pueden utilizar las técnicas denominadas *tiled-based streaming*, incluida en estándares como HEVC, que consisten en dividir o "trocear" el contenido en *tiles* (literalmente, baldosas) o pequeñas regiones. Sin embargo, las técnicas *tiled-based streaming* están fuera del alcance de esta Tesis Doctoral

Otra manera diferente de ahorrar ancho de banda se puede encontrar en [50], donde se presenta un reproductor de vídeo con soporte para la tecnología DASH [4] capaz de reproducir contenido codificado en *tiles* que selecciona la calidad más alta para las regiones dentro del FoV

y otras calidades inferiores para el resto de regiones. De esta manera, se hace un uso más eficiente del ancho de banda disponible. Para ello, se propone incluir la característica denominada MPEG-DASH *Spatial Representation Description* (DASH-SRD), definida en [51]. Esta característica se define en la segunda corrección del estándar MPEG-DASH Parte 1, 23009-1:2014, donde se define la metodología para transmitir únicamente regiones (*tiles*) específicas de un contenido multimedia. Esto se consigue al extender el fichero *Media Presentation Description* (MPD), encargado de almacenar en formato XML información de los segmentos que conforman el contenido, así como la relación temporal entre los segmentos y la información necesaria para que un cliente pueda escoger de entre las diferentes calidades disponibles; con una descripción con las relaciones espaciales entre las regiones o *tiles*.

2.11.2 CONTENIDO MULTISENSORIAL

Tradicionalmente, el contenido multimedia consumido ha involucrado exclusivamente flujos audiovisuales. Sin embargo, con el fin de estimular otros sentidos más allá de los tradicionales de la vista y el oído, ha surgido un nuevo concepto de consumo multimedia. Los contenidos multisensoriales permiten proporcionar efectos adicionales que estimulen al resto de sentidos durante el consumo de dicho contenido, con el objetivo de proporcionar al usuario final una experiencia de consumo multimedia más inmersiva y realista. Entre los muchos efectos sensoriales que se pueden incluir, se pueden citar los aromas, efectos ambientales (temperatura, efectos de iluminación, de niebla...), de vibración, etc. La Fig. 11 ilustra el concepto de mulsemmedia y, además, muestra, como ejemplo, algunos de los dispositivos que podrían estar involucrados en una experiencia de consumo mulsemmedia.

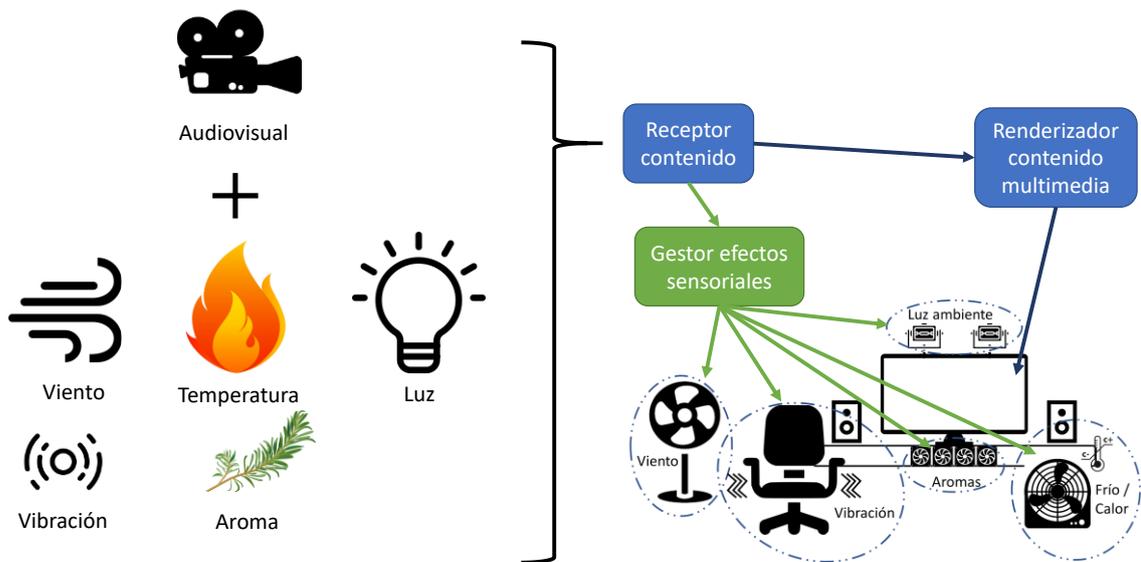


Fig. 11 Ejemplo de escenario de consumo de contenido mulsemmedia (adaptación de [52])

En el caso de experiencias de consumo mulsemmedia, se identifican una serie de aspectos importantes, como son la necesidad de reproducir de manera sincronizada todos los flujos involucrados en una experiencia de consumo mulsemmedia (tanto contenidos AV como la activación de ciertos efectos en momentos determinados), o la necesidad de disponer de herramientas de generación automática de este tipo de contenidos [52], [53], [54].

Respecto a su distribución, el estándar MPEG-V [55] propone una arquitectura y especifica la información necesaria para habilitar una interoperabilidad entre mundos virtuales y el mundo real. En este estándar, se consideran diferentes escenarios en los que los contenidos audiovisuales pueden ser asociados con efectos sensoriales que deben ser presentados en dispositivos específicos (por ejemplo, ventiladores, generadores de aromas, etc.). Además, en la Parte 3 de

este estándar se propone una sintaxis específica para poder anotar los efectos multisensoriales asociados a un contenido audiovisual, a través de un lenguaje denominado *Sensory Effect Description Language* (SEDL). Dicho lenguaje sigue un esquema XML y permite describir efectos sensoriales como luz, viendo, vibración, etc. Si bien los efectos sensoriales no forman parte del SEDL, están definidos en el *Sensory Effect Vocabulary* (SEV), para que cada aplicación o sistema que siga este estándar pueda integrar los efectos que considere oportunos. Finalmente, el fichero generado con el lenguaje SEDL se conoce como *Sensory Effect Metadata* (SEM), y puede asociarse a cualquier tipo de contenido multimedia, para así poder transmitir información multisensorial al consumidor de dicho contenido. El SEM se puede integrar en el flujo del contenido multimedia, o se puede adquirir a través de una red de transmisión diferente a la del contenido multimedia. La información descrita en el SEM se interpreta y mapea a comandos específicos de actuadores que son capaces de convertir la información en efectos sensoriales como viento (un ventilador), vibración (un motor), iluminación (luces inteligentes), etc.

2.12 CONCLUSIONES

En este capítulo se ha llevado a cabo una revisión de los conceptos, estándares y mecanismos que permiten el establecimiento de mecanismos de generación, transmisión y consumo sincronizado e inmersivo de contenido multimedia híbrido y multisensorial. Por un lado, se ha hecho una revisión de los estándares y middlewares relacionados con la transmisión y consumo de contenido híbrido. Por otro lado, se han presentado los diferentes tipos de sincronización, las estrategias de sincronización y los esquemas de control que pueden ser adoptados para alcanzar un nivel de sincronización satisfactorios. Finalmente, se ha expuesto las características más significativas de los contenidos de nueva generación (contenidos omnidireccional y multisensorial), así como las técnicas actuales para su codificación, transmisión y consumo.

Capítulo 3. TRABAJOS RELACIONADOS

En este capítulo se exponen trabajos relacionados con la temática de esta Tesis Doctoral, incluyendo trabajos relacionados con la sincronización híbrida entre dispositivos y/o destinatarios, con el caso de uso de la TV Social por ser el caso de uso escogido para evaluar la propuesta de esta Tesis Doctoral, así como con la generación, codificación y transmisión de nuevos contenidos, como los omnidireccionales o los multisensoriales, ya que se incluirán en los prototipos desarrollados en la Tesis Doctoral

3.1 SINCRONIZACIÓN HÍBRIDA

En este apartado se describen trabajos previos relacionados con soluciones de sincronización de la reproducción de contenidos híbridos, ya sea en un único dispositivo (terminal híbrido) o en varios dispositivos tanto locales (soluciones IDEs) como remotos (soluciones IDMS).

3.1.1 SINCRONIZACIÓN DE LA REPRODUCCIÓN DE CONTENIDOS HÍBRIDOS EN UN ÚNICO DISPOSITIVO (TERMINAL HÍBRIDO)

A continuación, se presentan trabajos relacionados con la sincronización de la reproducción de contenidos híbridos en un único dispositivo (terminal híbrido).

En [56], se propone un algoritmo capaz de sincronizar diferentes flujos MPEG-TS generados por el mismo proveedor de contenidos. Dicho algoritmo se basa en el control del proceso de reproducción de los flujos de audio y vídeo a partir de los valores de los campos PCR y PTS. Durante el inicio, se realiza un proceso de sincronización para asegurar que los flujos van a comenzar a presentarse en el mismo instante, retrasando los flujos con un PTS mayor hasta que todos los flujos involucrados coincidan en su PTS.

En [57], se describe un mecanismo para personalizar los servicios broadcast a través de Internet. En dicho trabajo, el broadcaster proporciona contenidos adicionales a través de Internet, los cuales están sincronizados con el contenido recibido a través de redes broadcast. Tanto el contenido broadcast como el transmitido a través de Internet, son enviados al mismo tiempo y el receptor de contenidos sincroniza los flujos a partir de los valores PCR y PTS. Generalmente, el contenido transmitido a través de Internet sufre mayores retardos, por lo que en dicho trabajo se implementa un buffer encargado de atrasar la reproducción del contenido broadcast y así conseguir un estado de sincronización en la reproducción de ambos tipos de contenido. Otro trabajo que utiliza los valores PCR y PTS para sincronizar flujos MPEG-TS híbridos se describe en [58]. En dicho trabajo se describe un caso de uso en el que se proporciona un flujo de audio alternativo vía broadband, adoptando el streaming basado en el protocolo RTP [2] sobre una red IP, con el fin de proporcionar más información (auditiva) para el caso de usuarios con dificultades o problemas de visión. Los contenidos son generados por diferentes entidades, por lo que se comparte una línea de tiempo global que se inserta como valores PTS en los flujos involucrados.

El trabajo en [59] propone una solución de transmisión de contenidos híbrida y de sincronización híbrida, denominada Advanced Transport Scheme (ATS). Dicha solución permite que los componentes o flujos multimedia puedan ser transmitidos de manera separada (o

multiplexados), a través de las diferentes redes disponibles. Sin embargo, no se evalúa la precisión de la sincronización alcanzada.

En los trabajos descritos en [60], [61] también se investiga la transmisión híbrida de contenidos. En particular, en [60] se propone una solución de sincronización híbrida entre un contenido audiovisual transmitido a través de la tecnología broadcast IPTV, utilizando para ello el protocolo RTP/RTCP [2], y un flujo de audio transmitido a través de DVB, siendo ambos transmitidos en flujos de transporte MPEG2-TS. En [61] se sincroniza un flujo broadcast MPEG2-TS audiovisual con un flujo de audio mp3 transmitido a través de radio por Internet. En dicho trabajo, ambos flujos proporcionan información acerca de un mismo partido de fútbol, lo cual permite a los usuarios elegir qué comentarista prefieren escuchar. La solución que se implementa en dicho trabajo consiste en eliminar el flujo de audio del flujo MPEG2-TS principal, transmitido a través de la tecnología DVB y generar un nuevo MPEG2-TS con un flujo de audio recibido a través de radio por Internet y relacionando ambos flujos con marcas de tiempo basadas en el protocolo NTP. En ambos trabajos ([60], [61]), se consigue un estado de sincronización satisfactorio gracias a la adopción de una línea temporal global y gracias a la implementación de dos procesos de sincronización diferentes: uno inicial y uno continuo. El proceso de sincronización inicial se basa en la utilización de marcas temporales basadas en NTP [21] para alinear la reproducción de los flujos multimedia involucrados. El proceso de sincronización continuo comienza tras finalizar el inicial, y consiste en la corrección periódica que puedan sufrir los relojes internos entre cada flujo multimedia. Si la transmisión se realiza a través de RTP/RTCP, se pueden adoptar las marcas de tiempo basadas en NTP de los mensajes Sender Reports de RTCP para sincronizar. Si la transmisión se basa en la tecnología DVB, se propone el uso de la tabla *Event Information Table* (EIT) dentro del flujo MPEG2-TS para contener las marcas de tiempo correspondientes a una línea temporal global.

En [62], se propone una solución que adopta el uso de relojes globales para la sincronización híbrida en un único dispositivo. En dicho trabajo se investigan diferentes escenarios, que se implementan y evalúan en una plataforma desarrollada a partir del framework GPAC [63]. Por un lado, se realiza una sincronización entre un flujo de audio transmitido por ondas FM y un flujo con contenido audiovisual MPEG2-TS a través de redes broadband. La sincronización se consigue a partir de insertar marcas de tiempo de un reloj global en la estructura Radio Data System (RDS) del canal FM y de la tabla Time and Date Table (TDT) dentro del MPEG2-TS. Por otro lado, también se implementa un caso en el que se sincronizan un flujo MPEG2-TS transmitido vía broadcast y un flujo MPEG2-TS transmitido vía broadband (empleando DASH [4]). En este caso, la información necesaria para sincronizar los contenidos se inserta en la tabla TDT de ambos flujos.

En [64], se propone un mecanismo de sincronización para contenidos híbridos (DVB y DASH), que se basa en la adopción de líneas de tiempo absolutas, comenzando desde el valor 0, correspondiente al inicio del programa. Respecto al flujo transmitido a través de la red broadcast (DVB), se adopta la especificación definida por la ETSI en [25], explicada en el Capítulo 2 de esta Tesis Doctoral. Respecto al contenido transmitido a través de la red broadband (DASH), al estar el contenido dividido en segmentos de igual duración, puede adaptarse de forma sencilla a la línea temporal absoluta en la parte del receptor.

En [65], se demuestra que es posible alcanzar un estado de sincronización híbrida del orden de milisegundos en un único dispositivo. Esto se consigue gracias a la posibilidad de combinar técnicas de sincronización que emplean tanto los valores de los campos PCR/PTS como el mecanismo definido por la ETSI en [25]. Ello permite evitar la problemática de emplear únicamente valores PCR/PTS (como por ejemplo, en [56] y [58]). En dicho trabajo, se utiliza el framework GStreamer [24] para demultiplexar el flujo MPEG2-TS y así poder analizar y comparar las líneas de tiempo globales insertadas en el flujo de transporte para conseguir un estado de sincronización aceptable. Para demostrar su utilidad, se reproduce contenido almacenado de manera local

(emulando contenido broadcast) junto con contenido broadband remoto obtenido mediante DASH. Un trabajo similar se describe en [66], también basado en el mecanismo de sincronización propuesto por la ETSI en [25] y en el uso de contenido DASH transmitido a través de redes broadband. En dicho trabajo se presentan diferentes herramientas compatibles con DASH para señalar contenido broadband adicional, así como la descripción de los contenidos transmitidos. Más recientemente, el mismo grupo de trabajo, presenta una plataforma de transmisión híbrida y de sincronización en [67]. En dicho trabajo se utiliza el framework GPAC [63] y permite insertar y/o extraer descriptores TEMI [26], reproducir vídeo codificado en HEVC, y utilizar DASH como tecnología de transmisión para contenido broadband.

3.1.2 SINCRONIZACIÓN DE LA REPRODUCCIÓN DE CONTENIDOS HÍBRIDOS EN VARIOS DISPOSITIVOS LOCALES

El consumo de contenidos híbridos relacionados en diferentes dispositivos locales de manera simultánea ha ido creciendo de manera significativa, y una muestra de ello es la cantidad de trabajos previos que abordan este tipo de entornos y la necesidad de establecer mecanismos de sincronización inter-dispositivo (IDES).

Respecto al descubrimiento de dispositivos y servicios, en [68] se realiza una revisión de los métodos y aproximaciones existentes, desde diferentes puntos de vista: a) las soluciones comerciales; b) los retos asociados al descubrimiento y asociación de los dispositivos en un entorno distribuido; y c) los retos que supone un entorno en el que la TV es el dispositivo en torno al cual giran el resto de dispositivos. En dicho trabajo, se presentan, además, soluciones implementadas en TV o Set-top boxes (STB) con el estándar HbbTV en las versiones 1.1 y 1.5, como, por ejemplo, diversas soluciones basadas en: i) información visual (por ejemplo, mostrar una URL con ID única para que el consumidor introduzca manualmente dicha URL en su dispositivos móvil o mostrar códigos QR); ii) basadas en patrones de audio (por ejemplo, superponiendo un patrón de sonido -transparente para el consumidor- que contenga una URL, aunque esta sea común para todos los consumidores, al ser transmitida esta información vía broadcast); o iii) basadas en Bump⁸ (actualmente obsoleto), una aplicación que permitía asociar la TV conectada y un smartphone cuando en la TV se mostraba un aviso y el consumidor "sacudía" el smartphone.

En relación con soluciones IDES de sincronización de la reproducción de contenidos híbridos en dispositivos locales, en [69] se adopta el mecanismo de la ETSI [25] combinado junto con mecanismos basados en los valores de los campos PCR/PTS, como una solución a la sincronización híbrida entre dispositivos en experiencias de TV con segundas pantallas. En dicho trabajo se afirma alcanzar un nivel de sincronización suficiente para proporcionar lip-sync⁹ (esto es, alrededor de 80ms) entre un flujo de audio recibido mediante RTP y multicast IP y un flujo de vídeo recibido mediante broadcast DVB-T. De manera similar, en [8] se incluye el desarrollo de una solución de streaming adaptativo adoptando el protocolo HTTP Live Streaming (HLS [3]) para la distribución de un contenido de vídeo dividido en regiones o tiles (*tiled-based streaming*). En dicho trabajo se utiliza el mecanismo propuesto por la ETSI en [25] para alcanzar un nivel de sincronización satisfactorio.

En el trabajo descrito en [70] se presenta un *framework* que incluye los mecanismos definidos en el estándar HbbTV para el descubrimiento y el establecimiento de las conexiones necesarias

⁸ <https://blog.bu.mp/post/71781606704/all-good-things>

⁹ El término lip-sync se refiere a sincronización labial, es decir, a la sincronización de los flujos de audio y vídeo, en vídeos en los que se visualizan movimientos labiales y que deben percibirse por el usuario de forma sincronizada con el audio asociado.

para proporcionar experiencias de consumo de TV con dispositivos/pantallas secundarias. Sin embargo, en dicho trabajo no se evalúa ningún tipo de solución de sincronización IDEs.

Respecto a la especificación DVB-CSS [71], que está reflejada en el estándar HbbTV y que permite alcanzar una sincronización entre un terminal híbrido (por ejemplo, una TV conectada) y dispositivos secundarios (CS), se ha implementado en los trabajos [72] y [73]. En particular, el trabajo en [72] emula la recepción de contenido broadcast a partir de un contenido MPEG2-TS almacenado localmente y presenta cuatro casos de uso: sincronización vídeo-vídeo, sincronización vídeo-audio, sincronización vídeo-aplicación web en el dispositivo secundario, y control activo o pasivo de la TV, dependiendo del uso de una aplicación web en el dispositivo secundario por parte del usuario. Por otro lado, en [73] un emulador de una TV, basado en Python, implementa los protocolos definidos en DVB-CSS¹⁰. En dicho trabajo, la TV muestra el contenido audiovisual principal mientras que los dispositivos secundarios presentan una interfaz web donde se muestran los subtítulos del contenido principal con diferentes niveles de asincronía (forzada). Adicionalmente, se implementan tres formas diferentes de interactuar con el usuario (pasiva - muestra texto sincronizado pero el consumidor no puede interactuar, solamente leer-, exploración -permite una interacción sencilla con el consumidor, mostrando información sobre actores/personajes-; y llamada a la acción -interactuando con el consumidor a través de preguntas y encuestas-), para evaluar el impacto en cuanto a la atracción y a la perceptibilidad del nivel de sincronización alcanzado. En dicho trabajo, se concluye que niveles de asincronía entre los -500ms (CS atrasado 0.5s) y los 1000ms (CS adelantado 1s) es muy probable que pasen desapercibidas para los usuarios en este tipo de experiencias con segundas pantallas. Además, los autores señalan que conforme aumenta el valor de la asincronía los consumidores sufren una mayor distracción por el contenido presentado en el dispositivo secundario.

En el trabajo expuesto en [74] se proporcionan resultados de una prueba piloto llevada a cabo durante unas elecciones en el País Vasco (España). El caso de uso implementado proporcionaba contenido híbrido multidispositivo durante un programa de TV en directo, incluyendo contenido complementario audiovisual (obtenido vía streaming basado en HLS), de redes sociales y de datos estadísticos y resultados de las elecciones (representados mediante gráficas y tablas en los CSs). En dicho trabajo, la sincronización no se consideró un aspecto crítico, ya que el proveedor de contenidos decidió adoptar una política de transmisión *best-effort*, para no añadir ningún tipo de retardo artificial en la transmisión broadcast de la señal de TV (contenido principal).

3.1.3 SINCRONIZACIÓN DE LA REPRODUCCIÓN DE CONTENIDOS HÍBRIDOS EN VARIOS DISPOSITIVOS REMOTOS

En [9], se analizan hasta 20 casos de uso diferentes en los que es requisito fundamental establecer mecanismos de sincronización IDMS. En [75], se indica que el rango aceptable de asincronía oscila entre los 15ms y los 500ms para escenarios IDMS, dependiendo del tipo de servicio. La siguiente Tabla muestra, de manera resumida, una serie de trabajos relacionados que presentan soluciones IDMS, incluyendo la tecnología en la que se apoyan dichas soluciones así como los casos de uso para los que se proponen (se puede encontrar una revisión más detallada de dichos artículos en los trabajos realizados por el grupo de I+D en el que se enmarca la presente Tesis Doctoral, [9], [76] y [77]).

¹⁰ <https://github.com/bbc/dvbcss-protocols>

Tabla 1 Clasificación de soluciones IDMS (adaptación de [9])

Referencia	Tecnología	Caso de uso
[78], [79], [80]	Protocolo propietario sobre UDP	Contenido bajo Demanda / Conferencias / Juegos / Mulsemedia / Entornos 3D Virtuales
[81], [82]	Protocolo propietario sobre UDP	Juegos [81], TV Social [82]
[77], [83], [84]	RTP/RTCP	TV Social / Conferencias / e-learning/ tele- vigilancia
[85]	DASH + P2P Overlay sobre UDP	(Web-based) TV Social (basada en web)
[86]	HTML5 + Websocket	TV Social (basada en web)

En [78] se propone la adopción del algoritmo llamado Virtual Time Rendering (VTR), que soporta la sincronización intra- e inter-media, pero extendiendo su aplicabilidad también para la sincronización IDMS para servicios de contenido bajo demanda. Desde dicho trabajo, el algoritmo VTR se ha adoptado extensamente para mejorar diferentes modelos de comunicación, casos de uso y escenarios como videoconferencias [79], y juegos en red y escenarios 3D virtuales y multisensoriales [80].

En [81] se propone una solución de sincronización IDMS para juegos en red, incluyendo dos algoritmos: el algoritmo "local lag", utilizado para compensar pequeños retardos a corto plazo; y el algoritmo "time warp", utilizado para corregir inconsistencias que puedan ocurrir de forma no controlada y que son inevitables. En [82], estos mismos algoritmos fueron adaptados para ser empleados en escenarios de TV Social.

En [83] y [84] se extienden las capacidades del protocolo RTP/RTCP [2], para ser compatible con mecanismos de sincronización IDMS, proponiendo dos soluciones bastante parecidas haciendo uso de dichas extensiones. La solución IDMS propuesta en [83] se evalúa en el propio trabajo y en [87], mientras que la solución propuesta en [84] se evalúa en [77].

En [85], se presenta una solución IDMS para contenido broadband transmitido mediante DASH. Dicha solución adopta un modelo de comunicación distribuida en la cual los clientes de una misma sesión construyen una capa de aplicación P2P sobre UDP a través de la cual intercambian la información temporal necesaria y negocian la referencia a la cual sincronizarse. Además, dicho trabajo también incluye una técnica de ajuste del proceso de reproducción basada en mecanismos Adaptive Media Playout (AMP), que considera la asincronía de los flujos de audio y vídeo para buscar el instante o periodo de tiempo más apropiado para realizar los ajustes en sus procesos de reproducción.

En [86] se presenta una plataforma basada en tecnologías web para el consumo distribuido de contenidos entre usuarios, habilitando además herramientas de interacción social. En dicha plataforma, se incluye una solución IDMS centralizada, basada en el esquema de control M/S, en la que el intercambio de mensajes con información temporal se realiza mediante la tecnología WS [32].

3.2 TV SOCIAL

Existen numerosos estudios previos que han analizado y señalado los beneficios que los escenarios de TV Social pueden aportar a los usuarios (por ejemplo, una sensación de estar en compañía o una mejora de las relaciones sociales). Sin embargo, la disponibilidad de un canal de comunicación entre los usuarios y de funcionalidades IDMS en dichos escenarios han sido

identificados en [88], [89] y en [90], [91], respectivamente, como requisitos clave para proporcionar dichos beneficios de forma satisfactoria.

Respecto a la disponibilidad de herramientas de comunicación mediante mensajes de texto (herramientas de chat), el hecho de escribir/leer el texto puede resultar una distracción para usuarios con pocas habilidades de escritura, los cuáles prefieren, generalmente, un canal de comunicación por voz [92]. Sin embargo, las comunicaciones mediante mensajes de texto es el canal preferido por usuarios con habilidades medias/altas de escritura, ya que esto les permite enviar mensajes mientras consumen el contenido audiovisual, sin ningún tipo de interferencias producidas por las voces del resto de usuarios. Además, esta comunicación suele producirse durante los periodos de silencio o cambio de escenas del contenido [88]. El estudio llevado a cabo en [93] también concluye que la modalidad de comunicación mediante mensajes de texto es la preferida por la mayoría de usuarios, ya que no es necesario responder de manera inmediata en la conversación, lo que provoca en los usuarios una sensación de que se requiere menos energía para interactuar con el resto de participantes. Por otro lado, en [91] también se afirma que tanto las herramientas de comunicación por voz, como por mensajes de texto, proporcionan una sensación de cercanía (*togetherness feeling*) con el resto de participantes.

En este contexto, diversos trabajos han estudiado en entornos experimentales controlados el efecto que tiene la existencia de diferentes niveles de diferencias de retardos sobre la QoE en escenarios de TV social, en los que los usuarios remotos se comunican mediante mensajes de texto y/o de voz durante la sesión [91], [94]. Resultados presentados en [91] indican que niveles de asincronía de hasta 1s podrían no ser perceptibles por los usuarios mientras se están comunicando a través de un canal de voz. Sin embargo, niveles por encima de 2s resultan molestos para los usuarios, independientemente de la vía de comunicación (texto o voz). En [94] se obtienen resultados similares.

En [95] se presenta un estudio realizado con el fin de determinar los niveles de asincronía tolerables entre contenidos multimedia interactivos consumidos por usuarios remotos. Tras realizar una evaluación subjetiva con usuarios, los resultados indican que niveles de asincronía alrededor de 400ms no producen un impacto negativo en la QoE percibida, mientras que asincronías a partir de 750ms ya son detectables por los usuarios y pueden afectar negativamente a la QoE.

Todos los trabajos anteriores relacionados con el caso de uso de TV Social consideran la integración del canal de comunicación mediante mensajes de texto en el mismo dispositivo donde se está reproduciendo el contenido principal. Sin embargo, en la actualidad se está generalizando el uso de dispositivos secundarios. En escenarios multidispositivo, se necesitarán soluciones de sincronización IDES, como las descritas en el apartado anterior, con el fin conseguir que la presentación de los contenidos reproducidos en cada dispositivo esté correctamente alineada a nivel temporal con respecto al punto de reproducción del contenido siendo presentado en el dispositivo principal.

Respecto al interés sobre este caso de uso específico de IDMS, en el trabajo descrito en [6], en el participó el autor de esta Tesis Doctoral, se presenta un estudio en el que participaron más de 1000 usuarios españoles, de los cuales alrededor del 80% mostró cierto interés en escenarios de TV Social, y más del 20% declaró haber participado en experiencias similares con anterioridad. En dicho trabajo se identifican los beneficios de este tipo de escenarios, tales como la sensación de compañía, la socialización y la atracción con los programas que presentan estas características. Por otro lado, más del 90% de los participantes declaró haber percibido retardos en dichos escenarios, y consideraron este problema como una barrera para poder consumir TV en una sesión compartida con usuarios remotos. La mayoría de los encuestados afirmaron que la disponibilidad de soluciones IDMS era requisito indispensable para poder disfrutar una

experiencia satisfactoria y que, además, todavía es necesario encontrar soluciones que funcionen con una mayor precisión que la disponible en la actualidad. Respecto a la modalidad del canal de interacción en escenarios de TV Social, un 54% afirmó preferir un canal de mensajería de texto.

3.3 VÍDEO OMNIDIRECCIONAL

En esta sección se exponen trabajos previos en los que se presentan diferentes mecanismos mediante los cuales poder transmitir de una manera más eficiente el contenido de vídeo omnidireccional, adoptando para ello los mecanismos y protocolos definidos en el Capítulo 2.

En [96] se adopta el mecanismo de *tiled-based streaming* definido en DASH-SRD para llevar a cabo una transmisión más eficiente de contenido omnidireccional. En dicho trabajo, se lleva a cabo una validación subjetiva con dispositivos como el HMD *Samsung Gear VR* y un smartphone *Samsung S7*. Según exponen en las conclusiones, se puede ahorrar hasta un 72% de ancho de banda al adoptar los mecanismos SRD definidos en [51].

El trabajo expuesto en [97] define tres mecanismos diferentes para la transmisión de contenido basado en tiles. Estos tres mecanismos son los siguientes, ordenados de peor a mejor eficiencia en cuanto al uso de ancho de banda: 1) *full delivery basic*, en el que todas las regiones (tiles) del contenido se entregan en la mejor calidad posible; 2) *full delivery advanced*, en el que todas las regiones son transmitidas pero aquellas que quedan fuera del FoV se descargan/solicitan con una menor calidad que la de las regiones incluidas dentro del FoV; y 3) *partial delivery*, en el que únicamente las regiones dentro del FoV se transmiten con la mayor calidad posible y el resto ni siquiera son descargadas/solicitadas.

Sin embargo, tal y como se expone en [98], el número de peticiones que realiza un cliente para recibir el contenido basado en tecnologías de transmisión adaptativa de contenidos basada en HTTP, como DASH, cuando adoptan mecanismos *tiled-based streaming*, es muy alto. Como ejemplo, los autores de dicho trabajo indican que existe una mayor carga de peticiones (de un 6% a un 22% más) cuando se consume contenido dividido en regiones y utilizando técnicas de *tiled-based streaming*, comparado con el caso en el que se consume el mismo contenido que no se ha dividido en regiones. Para solucionar esta sobrecarga, adoptan la funcionalidad *server-push* del protocolo HTTP/2. De esta manera, los clientes solo solicitan el contenido una vez y, a continuación, el resto de las regiones involucradas son transmitidas por parte del servidor, de manera automática. En dicho trabajo, también se hace un uso eficiente del ancho de banda, puesto que solamente se envían las regiones dentro del FoV con la mayor calidad posible.

Mientras que el *tiled-based streaming* apenas requiere cambios en la parte del servidor, implica una mayor carga en la parte correspondiente a los dispositivos de consumo, ya que son los encargados en determinar qué regiones descargar con una mayor (o menor) calidad, decodificarlas y unir las para poder renderizarlas/presentarlas en pantalla. Además, por un lado, la predicción de los movimientos del consumidor para determinar y anticiparse a su futuro posible FoV requiere de algoritmos complejos. Por otro lado, la decodificación de contenido multimedia omnidireccional transmitido vía *tiled-based streaming* requiere que los dispositivos móviles tengan una potencia de decodificación hardware significativamente alta, para poder reproducir el contenido de manera estable.

Por este motivo, se están desarrollando soluciones basadas en la nube (*Edge computing-based solutions*) que permiten eliminar cierta carga de procesamiento en los dispositivos de consumo, y llevarla a nodos de la red (*edge-nodes*) cercanos a los dispositivos, como por ejemplo, la decodificación de regiones con diferentes calidades según el FoV y su fusión en un único contenido que ya sí se transmite al usuario final.

En [99], se presenta un sistema que muestra contenido de vídeo omnidireccional con múltiples puntos de vista. En concreto, este trabajo presenta un contenido que permite ser navegado, a través de diferentes puntos de vista, habilitando así 6 grados de libertad (6DoF) al usuario. Para que exista un cambio rápido entre puntos de vista, los cambios deben poder ser anticipados por el sistema. Para conseguirlo, se consideran dos posibles estrategias para el cliente, una proactiva y una reactiva. Ambas estrategias tienen en cuenta tres variables: i) la distorsión y *bitrate* de cada segmento de vídeo codificado; ii) la trayectoria del usuario dentro del contenido; y iii) el ancho de banda medio disponible durante la descarga de cada segmento. Se concluye que el streaming de contenido dividido en *tiles*, mejora el rendimiento global del sistema, y que la estrategia proactiva es menos eficiente que la reactiva en términos de ancho de banda utilizado.

En [100], se presenta una plataforma web comercial a través de la cual se puede consumir contenido omnidireccional en directo. Durante una de las sesiones en directo, el sistema fue utilizado por un 10% de los consumidores habituales. Aunque dicha plataforma ofrecía el contenido mediante *tiled-based streaming*, y solamente las regiones correspondientes al FoV se transmitían en la mayor calidad posible, se evidenció que dichas regiones llegaban al cliente con un mayor retardo debido al mayor tamaño de sus segmentos.

3.4 MULSEMEDIA

Históricamente, mulsemedia ha sido un área ampliamente investigada [101]. Ya en un trabajo de 1999, [102], se presentan los resultados de un estudio subjetivo en esta área involucrando a 300 participantes. Se investigan los efectos provocados por el consumo de contenido mulsemedia (efectos olfativos), con el fin de determinar si dichos efectos sensoriales mejoran la sensación de presencia en escenarios de Realidad Virtual. Además, también se estudia cómo dichos efectos afectan a la memoria, comprobando si los participantes son capaces de recordar los objetos virtuales (dentro del escenario de RV). Los resultados en dicho trabajo demuestran que mejoran la sensación de presencia y que, además, ayudan a recordar mejor dichos objetos. De hecho, trabajos como [103] o [104] muestran que los contenidos multisensoriales permiten mejorar el rendimiento de aprendizaje, al estimular otro tipo de sentidos (por ejemplo, el olfato) como fuente complementaria de información. En la actualidad, el área mulsemedia continúa siendo objeto de investigación (por ejemplo, [16], [52], [101], [105], [106], [107], [108], [109], [110], [111]), desarrollándose numerosas aplicaciones para adoptar este tipo de contenido en entornos de RV o videojuegos, con el fin de estudiar de qué manera afecta a los usuarios y a la QoE percibida por ellos [112]. En [16], se enumeran los retos que acometer y problemas que resolver en esta área, como, por ejemplo, la necesidad de definir herramientas de autor para este tipo de contenidos. En [52], se presenta un estudio para entender el rol que podría adoptar el contenido mulsemedia (experimentan con luz ambiental, viento y efectos de vibración) en la mejora de la QoE de los usuarios en servicios de TV broadcast. En dicho trabajo, se concluye que el contenido mulsemedia puede: 1) mejorar la percepción de la calidad de vídeo consumida; 2) enmascarar (parcialmente) caídas eventuales en la calidad de la señal de vídeo recibida; y 3) proporcionar un mayor impacto en las emociones y sensaciones experimentadas. Sin embargo, también se afirma que existen numerosos retos por superar, como la generación (semi)automática de efectos sensoriales o el establecimiento de un umbral del nivel de asincronía tolerable entre estos efectos y los flujos audiovisuales.

En [105] se realiza una recopilación de trabajos relacionados con mulsemedia. Además, se enumeran diferentes áreas en las que el consumo de contenido mulsemedia puede mejorar la QoE percibida por los usuarios, como, por ejemplo, en áreas relacionadas con la interacción hombre-máquina (*Human-Computer Interaction* o HCI), telecomunicaciones, educación [104], *e-health*, *e-commerce*, publicidad, entretenimiento o turismo [113]. En [106], se definen una serie de

recomendaciones y pautas para diseñar y evaluar sistemas y aplicaciones con contenido mulsemmedia, centrando su foco de atención en el sentido del olfato. Además, en dicho trabajo también se enumeran los retos y dificultades clave que deben superarse para proporcionar al usuario una QoE satisfactoria en el ámbito mulsemmedia, como son la integración de efectos olfativos en un sistema multimedia, la sincronización, la estandarización, el desarrollo de dispositivos para el sistema olfativo, la intensidad y duración de los efectos olfativos, la aplicabilidad en distintas áreas (salud, educación, turismo...) y la entrega remota de los contenidos. En dicho trabajo, se indican los requerimientos necesarios para llevar a cabo una evaluación subjetiva de consumo de contenido mulsemmedia de forma satisfactoria: 1) la calidad mulsemmedia mínima respecto al diseño del entorno de consumo donde se va a consumir dicho contenido; 2) preparación/entrenamiento de los evaluadores (es decir, familiarización con este tipo de entornos/contenidos); 3) diseño experimental; y 4) diferentes características relacionadas con el olfato que deben tenerse en consideración, así como consejos a los evaluadores (por ejemplo, no beber café o evitar realizar ninguna evaluación si se está resfriado por la pérdida de capacidad olfativa). Al centrarse dicho trabajo en el sentido del olfato, recomiendan que durante una evaluación no se empleen más de 10 aromas diferentes, tal y como se define en [114].

En [115], se describe un trabajo que estudia la QoE percibida en contenidos mulsemmedia con efectos sensoriales olfativos incluyendo y omitiendo el componente auditivo. En concreto, se observa que cuando se presenta el contenido mulsemmedia sin audio, los instantes en los que no existe una sincronización entre la imagen y el efecto sensorial son menos molestos (y menos distinguibles) que cuando existe un componente auditivo.

Otro trabajo donde se realiza un estudio de diferentes dispositivos para el consumo de contenidos mulsemmedia es [107]. Se proporciona una guía para usuarios que quieran implementar su propio entorno mulsemmedia, bien en aplicaciones de escritorio (es decir, sobre un PC o portátil) o bien en escenarios 360 (a través de un HMD).

En [108], se propone el uso de contenido mulsemmedia para mejorar la QoE percibida en experiencias de consumo de contenido omnidireccional. En concreto, se expone que el contenido mulsemmedia mejora la QoE percibida en contenido multimedia tradicional y que, por tanto, se estudia incluir también este tipo de efectos multisensoriales para el consumo de contenido omnidireccional. En dicho trabajo se presenta un estudio de consumo de contenido 360 con y sin efectos multisensoriales, llegando a la conclusión de que, efectivamente, la QoE percibida mejora significativamente cuando se incluyen efectos multisensoriales.

Por otro lado, el trabajo presentado en [116] estudia efectos asociados a la QoE percibida, como, por ejemplo, un alteración en la frecuencia cardíaca durante escenas de miedo o suspense. En concreto, en dicho trabajo se monitoriza la frecuencia cardíaca y el movimiento de los ojos. Los autores de dicho trabajo concluyen que cuando se incluyen efectos sensoriales, concretamente olfativos, la QoE mejora significativamente comparado cuando solamente se presenta contenido audiovisual.

En [109] se explora el potencial del uso de mulsemmedia para mejorar la divulgación científica, con el fin de crear un mayor impacto en la atención de los usuarios. Se describen tres sesiones con comunicadores científicos donde se hace uso de dispositivos hápticos (percepción táctil a través de dispositivos de ultrasonido), donde los participantes señalan que estos dispositivos mejoran la experiencia compartida y que, además, facilitan la comunicación de conceptos científicos más complejos.

En [117], se realiza un estudio acerca de los factores humanos que influyen en la QoE percibida al consumir contenido mulsemmedia, señalando el sentido del olfato como uno de los más relevantes e influyentes en la QoE. Además, como ejemplo, en el artículo se señala que este tipo

de contenido mulsemmedia puede servir como herramienta para atraer a nuevas generaciones, mejor adaptadas a las tecnologías más recientes.

3.4.1 FRAMEWORKS, PROTOTIPOS Y PLATAFORMAS MULSEMEDIA

Respecto al desarrollo de frameworks o plataformas relacionadas con la generación y el consumo de contenido mulsemmedia, con el fin de mejorar la QoE del usuario o de llevar a cabo estudios subjetivos, existe un gran número de trabajos previos (por ejemplo, [118] [119] [120] [121] [122] [123] [124] [125] [126] [127] [128] [129] [130] [131]). En particular, en [122] se define una herramienta extremo-a-extremo (basada en el estándar *MPEG-V Part 3: Sensory Information* [55]) para la generación y el consumo de contenido mulsemmedia. En dicho trabajo, tras llevar a cabo una evaluación subjetiva, se concluye que: los contenidos mulsemmedia pueden llegar a ser molestos dependiendo del tipo de contenido (en noticias, por ejemplo); los efectos sensoriales tienen un gran impacto mejorando la percepción de la calidad de vídeo; las emociones son más intensas si hay efectos sensoriales involucrados en la experiencia de consumo.

Con el fin de señalar e integrar la información necesaria para poder reproducir contenido mulsemmedia, un framework denominado *Sensory Effect Metadata* (SEM) se propone en [118]. En dicho trabajo se propone el diseño de un controlador *Single Media Multiple Device* (SMMD), que se encarga de recibir el contenido mulsemmedia (incluyendo el audiovisual) y de enviar cada flujo al dispositivo correspondiente encargado de reproducirlo (el flujo audiovisual a una TV, la información de los olores asociados al contenido a un dispositivo emisor de aromas, información de temperatura a un generador de aire acondicionado, etc.). Todo ello es señalado a través del estándar MPEG-7 [132], el cual especifica una serie de metadatos para poder describir el contenido mulsemmedia. En dicho trabajo también se incluye la descripción de una herramienta que permite que usuarios no-expertos sean capaces de señalar información mulsemmedia en contenido originalmente audiovisual.

Otros trabajos como [119] y [120] presentan prototipos en los que se incluyen efectos como aromas, viento, luz o vibración. En [119] se propone una solución de sincronización que tiene en cuenta los retardos asociados al arranque de los dispositivos encargados de disparar los efectos sensoriales, así como el tiempo de transmisión de los datos asociados a los efectos. Por otro lado, en [120] se realiza una descripción detallada del hardware adoptado en un prototipo que incluye aromas (un vaporizador de aromas junto a un compresor de aire) y efectos de viento (un ventilador conectado a un PC y un compresor neumático -para emular "impactos" de viento-). En [123] se presenta un sistema mulsemmedia que incluye efectos sensoriales, como aroma, viento y háptico (táctil), con el fin de mejorar la experiencia de consumo audiovisual tradicional. Se lleva a cabo una evaluación subjetiva con contenido audiovisual de diferentes calidades (853x480 y 12801x720) junto con los efectos sensoriales mencionados. De dicho trabajo se concluye que el efecto sensorial preferido por los usuarios es el háptico, que se transmite al usuario a través de un chaleco preparado para tal fin. También se afirma que la inclusión de efectos sensoriales provoca un impacto positivo en la QoE percibida por el usuario, además de dotar a la experiencia de más realismo.

Por otro lado, el trabajo en [121] describe una herramienta para generar contenido mulsemmedia que incluye el reproductor incluido en el trabajo definido anteriormente (PlaySEM SER 2). Además, también incluye una herramienta de autor llamada *Sensory Effect Video Annotation* (Fig. 12). Dicha herramienta permite la generación y anotación de metadatos mulsemmedia asociados a un contenido audiovisual, que quedan anotados en un fichero XML, siguiendo la sintaxis definidas en el estándar MPEG-V [55]. Por otro lado, en dicho trabajo también se presenta un simulador denominado *Sensory Effect Simulator* (SESim, Fig. 13), que permite a los usuarios simular y reproducir un contenido mulsemmedia en un PC, donde los efectos sensoriales quedan emulados gráficamente a través de dispositivos virtuales (por ejemplo, con una imagen de un ventilador que

se activa con una determinada intensidad, o luces ambientales que muestran un color rgb específico para cada instante).

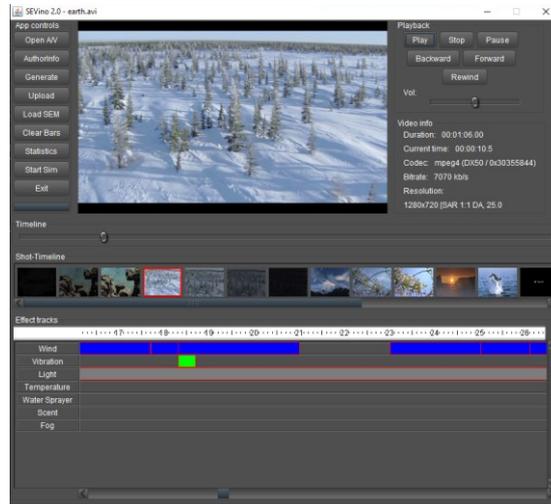


Fig. 12 SEVino 2.0 tool [121]

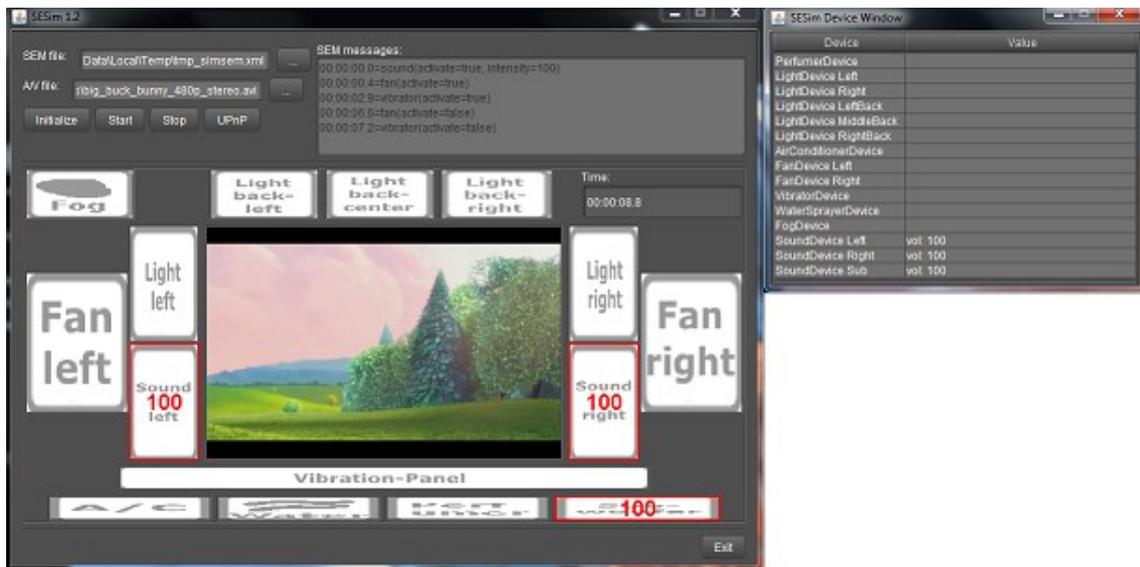


Fig. 13 Interfaz del simulador SESim

En el trabajo presentado en [124] se analiza la QoE percibida por usuarios en un entorno de consumo de contenido mulsemmedia basado en tecnología web. Concretamente, se estudia cómo influyen en la experiencia de consumo los efectos de luz ambiental, así como el impacto asociado a diferentes configuraciones para calcular el color que debe recrearse. Al igual que otros trabajos descritos con anterioridad, adopta el formato SEM para señalar y describir los efectos multisensoriales. Los resultados de una evaluación subjetiva demuestran que, en efecto, la inclusión de efectos multisensoriales mejoran la QoE percibida por los usuarios, conclusión a la que también llegan los mismos autores en el trabajo previo [133]. Adicionalmente, en dicho trabajo se describe un plugin para el navegador Firefox, capaz de interpretar datos SEM, aunque actualmente no se encuentra disponible.

En [126], se presenta un framework (denominado ADAMS) que también adopta el estándar MPEG-7 [132] para la generación y consumo de contenido mulsemmedia. Dicho trabajo adopta mecanismos de sincronización gruesa y fina (del inglés *coarse-* y *fine-grained*), basándose en la adaptabilidad del flujo mulsemmedia y en un establecimiento de prioridades en los paquetes del

flujo mulsemmedia, respectivamente. En este trabajo se expone que ambos mecanismos de sincronización mejoran la QoE percibida por los usuarios en diferentes condiciones de red, demostrando cómo dicho framework es capaz de minimizar el impacto en la QoE percibida.

De manera similar, en [127] se presenta otra herramienta de autor basada en el estándar MPEG-V. Dicha herramienta se divide en dos partes: la encargada de la generación de los metadatos asociados a los efectos sensoriales y la encargada de la simulación de los efectos sensoriales junto con el contenido audiovisual en un entorno virtual 3D. En particular, para la generación de los metadatos para los efectos sensoriales, se adopta el lenguaje de descripción de efectos sensoriales (*Sensory Effect Description Language* o SEDL), definido en [134].

En [135] se implementa un prototipo para emular comida y bebida en un escenario virtual, a través de dispositivos electrónicos. Estos dispositivos permiten al usuario probar ingredientes a través de electrodos que estimulan el sentido del gusto (en dos dispositivos diferentes, llamados Spoon+ -una cuchara- y Bottle+ -una botella-). Además, estos dispositivos cuentan con LEDs que permiten emular el color o las tonalidades de los ingredientes que están siendo "probados". Para demostrar la efectividad de estos dispositivos, se llevan a cabo dos experimentos, en los que se debe evaluar la intensidad de sensaciones de sabores como el salado, el agrio o el amargo. Los resultados demuestran que los usuarios distinguen de una manera más precisa el salado y el agrio.

El trabajo en [128] también presenta un framework para codificar, transmitir y consumir este tipo de contenidos de una manera totalmente inmersiva. Los mismos autores, presentan una descripción en detalle de dicho framework, denominado *PlaySEM SER 2*, en [130]. Este framework es una evolución de la plataforma *PlaySEM Sensory Effects Render (SER)*, presentada en [125]. En concreto, *PlaySEM SER 2* soporta múltiples protocolos de conectividad y comunicación, diferentes estándares (como MPEG-V [55]) y permite la integración de nuevas tecnologías dentro de su arquitectura. Este framework permite la reproducción de efectos sensoriales tales como aroma, vibración, iluminación o viento. En [128] se presentan y evalúan diferentes casos de uso para demostrar la reusabilidad de estos efectos y su adaptabilidad a diferentes escenarios, como, por ejemplo, un videoclip enriquecido con iluminación ambiental, aromas, vibración y viento; o un sistema únicamente con aromas.

En [131], se presenta un prototipo conceptual donde se integra contenido mulsemmedia con el fin de mejorar la QoE percibida durante el consumo de contenido omnidireccional. En dicho trabajo, se indica que añadir efectos multisensoriales permite reducir la resolución del contenido audiovisual omnidireccional, manteniendo una QoE percibida por el usuario en niveles aceptables.

3.4.2 TRANSMISIÓN DE CONTENIDOS MULSEMEDIA A TRAVÉS DE REDES BROADCAST

A pesar de ser técnicamente posible, no existe un gran número de trabajos previos que adopten las tecnologías broadcast para la emisión de contenido mulsemmedia. En [136] se propone un sistema basado en tecnologías broadcast en el que se incluye información relacionada con efectos sensoriales asociados al contenido audiovisual. Para tal propósito, se definen y señalizan los requisitos multisensoriales en un fichero XML que se integra en el flujo MPEG-TS durante su codificación a través de la plataforma que exponen en dicho artículo. De manera similar, en [137] se define un framework para transmitir contenido mulsemmedia, encapsulado en paquetes MPEG2-TS, a través de redes broadcast. En dicho trabajo se adopta el estándar MPEG-V [55], pero no se lleva a cabo ningún tipo de evaluación. Sin embargo, en ninguno de los dos artículos detallan el procedimiento de integración y transmisión del contenido mulsemmedia, más allá de indicar el tipo de formato en el que lo encapsulan (MPEG2-TS) y que el tipo de transmisión es broadcast.

Un entorno mulsemmedia basado en IPTV se presenta en [138]. En dicho trabajo el SEM se registra en un fichero XML, el cual se integra en el flujo MPEG-TS. Dicho flujo se transmite hasta el receptor del hogar, donde los dispositivos involucrados para los distintos efectos sensoriales

son capaces de decodificar y analizar los datos relacionados. Sin embargo, en dicho trabajo no se detalla cómo se integra en el flujo MPEG2-TS.

En los trabajos [139] y [140] se implementa un prototipo, capaz de presentar contenido mulsemmedia en un escenario de hogar inteligente (esto es, con Smart TV y dispositivos IoT capaces de generar los efectos sensoriales). En dicho trabajo se adoptan dispositivos comunes como el aparato de aire acondicionado o los dispositivos de luz ambiental como generadores de los efectos sensoriales. En particular, emplean la herramienta de reproducción de contenido mulsemmedia presentada en [125], con contenido almacenado en local. Los resultados de evaluaciones subjetivas muestran la viabilidad y utilidad de los servicios de TV multisensoriales. En esta Tesis Doctoral, a diferencia de estos dos trabajos, se emplea contenido broadcast transmitido a través de la tecnología DVB-T.

3.4.3 SINCRONIZACIÓN DE CONTENIDO MULSEMEDIA

En [141] el contenido mulsemmedia se separa en diferentes flujos, dependiendo del sentido que esté involucrado (como por ejemplo, calor, viento, aromas, luz o vibración). El algoritmo de sincronización que se propone en dicho trabajo está basado en el uso del protocolo RTP/RTCP y consiste en un esquema M/S donde el dispositivo encargado de presentar el contenido audiovisual adopta el rol de máster. En dicho trabajo se lleva a cabo una evaluación donde se activa y desactiva el algoritmo de sincronización con el objetivo de comprobar los valores de asincronía entre contenido audiovisual y multisensorial cuando el algoritmo está activo y cuando no lo está. Se concluye que, para proporcionar experiencias de consumo mulsemmedia satisfactorias, se debe integrar un mecanismo de sincronización, puesto que el valor de asincronía cuando el algoritmo no está activo supera los 100ms y cuando está activo está en valores alrededor de los 30ms.

En [142] se estudia la QoE percibida cuando dos flujos de contenido multisensorial olfativo se sincronizan con un contenido multimedia convencional (audiovisual). Este estudio se lleva a cabo para determinar si el nivel de sincronización alcanzado entre los efectos multisensoriales y el contenido audiovisual, así como el jitter, pueden afectar al QoE percibido por los usuarios. En dicho trabajo se llega a la conclusión de que los efectos multisensoriales olfativos pueden considerarse como sincronizados si su nivel de sincronización está comprendido entre los valores [-10, 15] segundos (adelantados o retrasados, respectivamente), y como desincronizados si el valor de sincronización alcanzado se encuentra fuera del rango.

Posteriormente, en [143] se expone un nuevo estudio sobre la relación temporal entre el contenido audiovisual y efectos sensoriales olfativos. En dicho trabajo se lleva a cabo una evaluación para analizar la habilidad de los usuarios para detectar asincronías y, por tanto, conocer cómo afecta a la QoE. Se concluye que el contenido mulsemmedia (olfativo) se acepta como satisfactoriamente sincronizado si está entre [-5, 10]¹¹ segundos del instante del contenido audiovisual relacionado.

3.5 CONCLUSIONES

Por un lado, en relación con los mecanismos de sincronización IDEs, en la Tesis Doctoral se adopta el estándar HbbTV como la base para la solución propuesta en la misma, debido a su reciente implantación y su relevancia en el nuevo paradigma de consumo de contenidos de TV híbridos. Sin embargo, dicho estándar no proporciona soluciones específicas a requerimientos

¹¹ Un valor negativo significa que el efecto sensorial ocurre antes que el evento audiovisual y un valor positivo significa que dicho efecto ocurre a posteriori.

para poder implantar, de manera satisfactoria, un mecanismo de sincronización entre dispositivos (IDES). Las necesidades más destacables son:

- Mecanismos de señalización para descubrir, asociar y describir el contenido relacionado disponible.
- Mecanismos de interacción y coordinación entre los dispositivos secundarios disponibles.
- Soluciones de sincronización completamente desarrolladas (incluyendo protocolos, algoritmos y técnicas de ajuste del proceso de reproducción) para alinear de manera precisa el contenido multimedia que está siendo consumido en uno o varios dispositivos, ya sea el mismo contenido o contenidos relacionados. En particular, esta necesidad es particularmente relevante y compleja, debido a la variabilidad de los retardos cuando se transmiten diferentes contenidos a través de diferentes redes, incluso desde diferentes proveedores de contenido, y se consumen en un único dispositivo (escenario mono-dispositivo, como, por ejemplo, una TV conectada) o en más de un dispositivo (escenario multidispositivo), con todo lo que esto supone (especificaciones, capacidades y conectividad heterogéneas).

Esta Tesis Doctoral aporta las siguientes contribuciones relacionadas con escenarios mono- y multidispositivo (IDES): i) especificación y diseño de necesidades clave para alcanzar un nivel de sincronización IDES satisfactorio, como son la señalización de contenido relacionado; ii) la interacción y coordinación entre dispositivos principales y secundarios; y iii) mecanismos específicos de sincronización.

Por otro lado, respecto a los mecanismos de sincronización IDMS, actualmente siguen existiendo una serie de limitaciones:

- La mayoría están implementados con protocolos propietarios (por ejemplo, [78], [79], [80], [81], [82]) o con técnicas propietarias. Sin embargo, tal y como se expone en [34], [144], la disponibilidad de soluciones estándar (o compatibles con estándares) proporcionará grandes beneficios.
- Están diseñados exclusivamente para entornos cuya tecnología de transmisión es broadband [77], [78], [79], [80], [81], [82], [83], [84], [85], [145]. Sin embargo, la tecnología de transmisión broadcast puede seguir considerándose la más apropiada para llegar a una mayor audiencia [146]. Estudios previos han demostrado que existen diferencias de retardos significativas en escenarios broadcast [147], [148], por lo que también son necesarias soluciones IDMS cuando el contenido se transmite a través de redes broadcast.
- Todas las soluciones IDMS propuestas hasta la fecha funcionan únicamente para un solo tipo de tecnología de transmisión (por ejemplo, [77], [78], [79], [80], [81], [82], [83], [84], [85], [145]). No combinan más de una tecnología de transmisión para una experiencia de consumo compartida. Sin embargo, resulta lógico asumir que un escenario así es realista en la actualidad, donde cada uno de los participantes de una sesión compartida puede obtener el contenido multimedia a través de distintas redes de transmisión.
- Las soluciones actuales asumen que el contenido compartido es exactamente el mismo para todos y cada uno de los usuarios en una sesión (por ejemplo, [77], [78], [79], [80], [81], [82], [83], [84], [85], [145]), independientemente del dispositivo que se esté utilizando para consumirlo.
- Actualmente, no existen soluciones IDMS implementadas y/o validadas en entornos realistas y a partir de prototipos cercanos a una versión comercial, excepto las implementadas por el autor de la Tesis Doctoral y el grupo de investigación del que forma parte.

La solución de sincronización IDMS que se propone en esta Tesis Doctoral supera todas las limitaciones anteriores. Primero, se basa en protocolos de dominio público. Segundo, el mecanismo de sincronización propuesto soporta tanto contenido transmitido por tecnologías broadcast como broadband. Ambas tecnologías pueden utilizarse en una sesión compartida, de manera que un usuario pueda recibir el contenido a través de la red broadcast y otro usuario a través de la red broadband. Tercero, se puede consumir más de un contenido en la misma sesión compartida. Como ejemplo, un usuario puede consumir un contenido que ha recibido a través de la red de transmisión broadcast, y otro usuario puede estar consumiendo otro contenido del mismo evento (por ejemplo, otro punto de vista distinto), que haya podido recibir a través de la red broadband. Además, el mecanismo propuesto se ha validado en un entorno cercano a un escenario real, puesto que uno de los objetivos principales es mejorar los servicios de consumo de contenidos híbridos actuales.

Por lo tanto, esta Tesis Doctoral aporta las siguientes contribuciones relacionadas con escenarios multidestinatario (IDMS): i) especificación y diseño de necesidades clave para alcanzar un nivel de sincronización IDMS satisfactorio, como son la señalización sesiones de consumo de contenido audiovisual disponibles; ii) la interacción y coordinación entre dispositivos principales en cada destinatario; iii) mecanismos específicos de sincronización; y iv) canales de comunicación sincronizados entre los usuarios de una misma sesión de consumo de contenido audiovisual.

Esta Tesis Doctoral presenta un escenario de evaluación basado en el caso de uso de la TV Social, en el que se trata de dar una solución a algunos de los retos que todavía existen para proporcionar experiencias de consumo compartidas, tales como la sincronización multimedia, la escalabilidad, la QoE, la integración de redes sociales o problemas de privacidad que pueden derivarse de este tipo de escenarios. En particular, esta Tesis Doctoral se centra, principalmente, en proporcionar unos niveles de sincronización IDMS/IDES satisfactorios, así como unos canales de interacción entre usuarios de una sesión de consumo compartida.

Respecto a la generación y consumo de contenido mulsemmedia, las principales limitaciones guardan relación con la transmisión de este tipo de contenidos a través de redes broadcast, por lo que las principales contribuciones en esta área en esta Tesis Doctoral son: i) el diseño e implementación de una solución capaz de generar, codificar (los metadatos asociados) y presentar contenido mulsemmedia relacionado con contenidos recibidos a través de redes broadcast y broadband; ii) un mecanismo que permite señalar la existencia de efectos sensoriales de forma integrada dentro del flujo de transporte MPEG2-TS; iii) el diseño de un controlador de efectos multisensoriales, capaz de analizar y dar soporte a cualquier efecto sensorial (siempre y cuando se disponga de los materiales y dispositivos encargados de generar dichos efectos); iv) diseño de una interfaz (API) para la comunicación entre los dispositivos de consumo y el controlador mulsemmedia; v) diseño e implementación de un mecanismo IDMS para sincronizar la reproducción del contenido audiovisual y la generación de los efectos sensoriales; y vi) la integración del soporte para contenido mulsemmedia en la plataforma extremo-a-extremo del grupo de I+D en el que se enmarca esta Tesis Doctoral, adaptada para los casos de uso de los escenarios de evaluación.

Capítulo 4. SOLUCIONES DE SINCRONIZACIÓN MULTIMEDIA Y MULTISENSORIAL HÍBRIDA IDES E IDMS

En este capítulo se presentan el conjunto de mecanismos de sincronización híbrida propuesto para posibilitar nuevas experiencias televisivas enriquecidas, personalizadas, inmersivas y compartidas.

4.1 INTRODUCCIÓN

Las soluciones de sincronización híbridas propuestas en la Tesis Doctoral son válidas para los siguientes escenarios:

- Escenario IDES: consiste en un entorno de consumo local de contenidos multimedia híbridos (en 2D o 360) en el que existe un dispositivo principal o MS y uno o varios dispositivos secundarios o CS, todos ellos físicamente próximos (por ejemplo, en una misma red de área local).
- Escenario IDMS: consiste en un entorno de consumo distribuido de contenido multimedia (en 2D o 360) en el que existen dos o más dispositivos principales o MS, cada uno de ellos separados geográficamente (por ejemplo, en diferentes hogares).
- Escenario IDES + IDMS: consiste en un entorno distribuido de consumo de contenidos multimedia híbridos (en 2D o 360) en el que existen dos o más dispositivos principales o MS distribuidos y uno o varios dispositivos secundarios o CS asociados a cada uno de ellos, estando organizados en grupos separados geográficamente. Se trata de una combinación de los dos primeros escenarios.
- Escenario mulsemmedia: consiste en un entorno que por sus características es compatible con cualquiera de los escenarios anteriores que, además, incluye efectos multisensoriales en la experiencia de consumo (como, por ejemplo, contenido 2D en el MS, contenido 360 en el CS y dispositivos sensoriales generadores de efectos, como aromas o viento).

Para que la solución de sincronización de los procesos de reproducción (ya sea involucrando un único dispositivo, o varios dispositivos locales o remotos) funcione correctamente, se debe cumplir con una serie de requisitos previos:

- 1) Los contenidos híbridos deben estar relacionados de alguna manera, con el fin, por ejemplo, de poder localizar contenido complementario al contenido principal.
- 2) Los procesos, dispositivos o destinatarios (según escenario) deben descubrirse y deben establecer algún tipo de canal de comunicación bidireccional entre ellos que les permita comunicarse entre sí.

- 3) Los procesos, dispositivos o destinatarios involucrados en la misma sesión de consumo deben soportar y reproducir contenido recibido a través de la red broadcast y también a través de la red broadband.
- 4) Los procesos de reproducción involucrados deben soportar mecanismos de ajuste del estado de reproducción con el fin de alcanzar un nivel de sincronización satisfactorio.

En las siguientes secciones se detallan los estándares, protocolos y tecnologías que se han adoptado en esta Tesis Doctoral para poder cumplir con estos requisitos previos.

4.2 TECNOLOGÍAS, ESTÁNDARES Y PROTOCOLOS ADOPTADOS

Tal y como se ha expuesto en el Capítulo 2, en la actualidad, el estándar HbbTV permite a los proveedores de contenido proporcionar contenido multimedia a través de redes broadband, mientras transmiten el contenido principal a través de redes broadcast (por ejemplo, DVB o IPTV). Es por ello por lo que en esta Tesis Doctoral se han adoptado los requisitos y especificaciones definidos en dicho estándar para implementar diferentes soluciones de sincronización entre dispositivos y/o entre destinatarios. Aunque este último escenario no está contemplado en dicho estándar, las tecnologías adoptadas en el mismo permiten la implementación de este tipo de funcionalidades.

4.2.1 TRANSMISIÓN HÍBRIDA DE CONTENIDOS

En esta Tesis Doctoral, con el fin de proporcionar una solución de sincronización que pueda ser aplicada en entornos reales, para la transmisión broadcast de contenido se contempla el uso de la tecnología DVB. En el caso de la transmisión a través de redes broadband, las tecnologías contempladas son MPEG-DASH (adoptada por la especificación HbbTV actual) y, adicionalmente, las tecnologías HLS y RTP/RTCP, con el fin de aumentar el soporte y compatibilidad con todo tipo de dispositivo secundarios (Fig. 14).



Fig. 14 Tecnologías de transmisión adoptadas en esta Tesis Doctoral

4.2.2 SEÑALIZACIÓN DE EFECTOS MULTISENSORIALES

Respecto a la señalización del contenido mulsemmedia, el estándar adoptado en la Tesis Doctoral es el MPEG-V [55], el cual incluye la estructura y sintaxis descrita en la sección 2.11.2 de esta Tesis Doctoral, denominada *Sensory Effect Description Language (SEDL)*, definido en [134] y que permite describir efectos sensoriales en un fichero XML, así como atributos tales como intensidad, duración, prioridad o parámetros específicos de cada efecto (por ejemplo, el tipo de aroma). En la Fig. 15 se muestra un ejemplo de un fichero XML con las etiquetas que se emplean para la definición de los efectos multisensoriales.

```

<?xml version="1.0" encoding="UTF-8"?>
<sedl:SEM [...]>
  <sedl:Effect xsi:type="sev:LightType" autoExtraction="both" activate="true"
    si:anchorElement="true" si:pts="0"/>
  <sedl:Effect xsi:type="sev:WindType" intensity-value="100.0" intensity-range="0.0 100.0"
    activate="true" si:anchorElement="true" si:pts="1620000"/>
  <sedl:Effect xsi:type="sev:WindType" activate="false" si:anchorElement="true" si:pts="2160000"/>
  <sedl:Effect xsi:type="sev:ScentType" intensity-value="100.0" intensity-range="0.0 100.0"
    activate="true" sev:scent="smoke" si:anchorElement="true" si:pts="4149000"/>
  <sedl:Effect xsi:type="sev:ScentType" activate="false" sev:scent="smoke"
    si:anchorElement="true" si:pts="4689000"/>
  <sedl:Effect xsi:type="sev:VibrationType" intensity-value="25.0" intensity-range="0.0 100.0"
    activate="true" si:anchorElement="true" si:pts="14229000"/>
  <sedl:Effect xsi:type="sev:VibrationType" activate="false" si:anchorElement="true" si:pts="15579000"/>
  <sedl:Effect xsi:type="sev:LightType" activate="false" si:anchorElement="true" si:pts="28629000"/>
</sedl:SEM>

```

Fig. 15 ejemplo de fichero SEM

Tal y como se observa en la imagen, cada efecto se identifica con una etiqueta *sedl:Effect* y contiene parámetros tales como *xsi:type*, *intensity-value*, *intensity-range*, *activate* y *pts*; los cuales permiten definir la naturaleza del efecto (por ejemplo, aroma, viento, luz ambiental o vibración), la intensidad a la que se debe presentar el efecto, el rango de intensidad para dicho efecto, si debe activarse o desactivarse y cuándo se debe presentar el efecto, respectivamente.

4.2.3 DESCUBRIMIENTO Y COMUNICACIÓN ENTRE DISPOSITIVOS

Tal y como se ha comentado en capítulos anteriores, en un escenario multidispositivo, se requiere que los dispositivos involucrados estén vinculados, de alguna manera, entre sí y dispongan de un canal de comunicación bidireccional para poder intercambiar información útil para sincronizar sus procesos de reproducción. Para poder vincularlos, es necesario un mecanismo de descubrimiento previo. Para tal fin, en la presente Tesis Doctoral se ha adoptado el protocolo DIAL [30] para que los dispositivos involucrados en un entorno local (ejecutando las aplicaciones MS y CS) puedan descubrirse y establecer dicho canal de comunicación entre ellos. A través del protocolo SSDP un dispositivo puede descubrir al resto de dispositivos que estén conectados en su misma red de área local. Las solicitudes de descubrimiento pueden enviarse tanto por parte del dispositivo principal como de los secundarios. Una vez los dispositivos se 'conocen', se puede establecer un canal bidireccional de comunicación directo entre ellos (canal *App2App*) mediante tecnología Websocket [2], que deberá ser soportada por todos los dispositivos involucrados. La Fig. 16 muestra los componentes y funcionalidades de comunicación involucrados en un entorno local multidispositivo, como el de algunos de los escenarios considerados en la Tesis Doctoral. Una vez creado dicho canal de comunicación, los dispositivos pueden intercambiar la información necesaria según el tipo de escenario en el que se encuentren (por ejemplo, el MS puede enviar al CS información o metadatos del contenido relacionado, como, por ejemplo, cómo poder acceder al mismo a través de redes broadband).

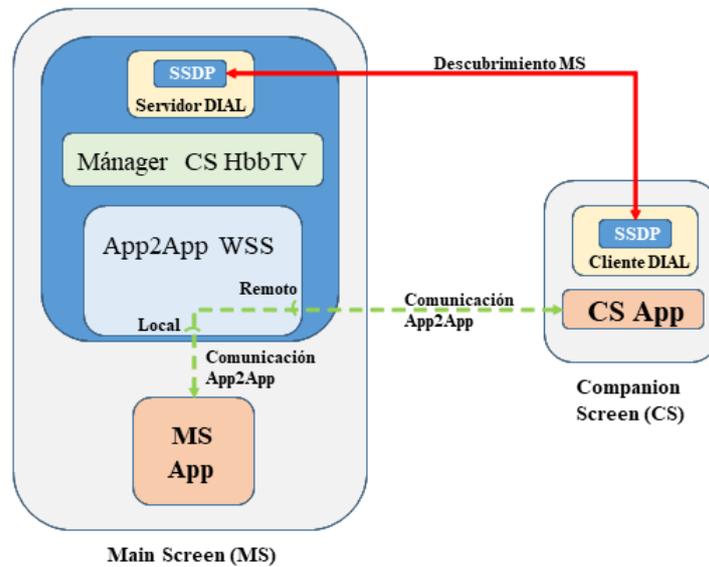


Fig. 16 Componentes y funcionalidades de comunicación involucrados en un entorno local multidispositivo basado en HbbTV y DIAL

4.3 PREPARACIÓN DEL CONTENIDO

En este subapartado se describe el método propuesto para la generación y señalización del contenido.

4.3.1 INSERCIÓN DE MARCAS TEMPORALES EN EL CONTENIDO PRINCIPAL

Tal y como se ha expuesto en el apartado anterior, el contenido broadcast se transmite a través de DVB, por lo que el contenido audiovisual se encuentra encapsulado en paquetes MPEG2-TS. De esta manera, se puede adoptar el mecanismo TEMI [26] descrito en el capítulo 2, para incluir información útil para la sincronización. En esta Tesis Doctoral se propone utilizar el descriptor TEMI *timeline* para insertar marcas temporales (por ejemplo, de un reloj global) en el contenido para poder alcanzar un nivel de sincronización satisfactorio, ya sea entre destinatarios o entre dispositivos (situaciones en las que se requiere de mecanismos IDMS e IDES, respectivamente).

4.3.2 SEÑALIZACIÓN DE LA EXISTENCIA DE CONTENIDOS COMPLEMENTARIOS VÍA REDES BROADBAND

Con el fin de señalar la disponibilidad de contenidos broadband complementarios dentro del flujo broadcast, en esta Tesis Doctoral también se adopta la solución denominada TEMI. Para ello, se propone utilizar el descriptor TEMI *location* para indicar la ubicación (esto es, la URL) de un fichero con formato XML denominado *Related Media Contents File* (RMCF), donde se definen (metadatos) y señalizan todos los contenidos (complementarios y relacionados con el contenido broadcast o principal) disponibles a través de redes broadband, además de otro tipo de información necesaria (como por ejemplo, la dirección URL de un servidor de tiempo global – como, por ejemplo, un servidor NTP–, para poder sincronizar los procesos de reproducción involucrados, o la dirección URL de un servidor de chat privado para que un grupo de amigos puedan comunicarse durante una sesión compartida de consumo de contenidos multimedia).

Una vez el dispositivo principal (por ejemplo, una TV conectada) haya recibido el contenido principal vía broadcast, debe ser capaz de identificar, extraer y analizar los descriptors TEMI, con

el fin de descubrir (y notificar a los dispositivos secundarios involucrados, si existen) el contenido complementario, disponible a través de redes broadband. Por este motivo, en esta Tesis Doctoral, se han definido una serie de etiquetas y propiedades siguiendo el formato XML, con el fin de que dicho fichero pueda ser analizado con facilidad. En la Tabla 2 se exponen todas las etiquetas definidas, incluyendo aquellas específicas para escenarios IDES, IDMS y mulsemedia, que se comentarán más adelante en este capítulo.

Tabla 2 Etiquetas y propiedades definidas en el fichero RMCF

Etiqueta	Escenario	Propiedad	Descripción
MEDIA	IDES, IDMS, Mulsemedia		Especifica los metadatos necesarios para cualquier contenido AV o mulsemedia disponible.
		<i>id</i>	Valor de identificación único para el elemento.
		<i>media_type</i>	Tipo de contenido: 'AV' (contenido tradicional), '360AV' (contenido omnidireccional), 'SEM' (contenido mulsemedia).
		<i>media_format</i>	Información del formato o codificación (por ejemplo, H.264).
		<i>metadata</i>	Breve descripción.
		<i>temi_init</i>	Marca de tiempo global absoluta cuando se inició la generación del contenido.
Source (dentro de la etiqueta MEDIA)	IDES, IDMS, Mulsemedia		Permite especificar fuentes alternativas para el mismo contenido.
		<i>protocol</i>	Protocolo de streaming utilizado (por ejemplo, DASH).
		<i>projection</i>	Tipo de proyección utilizado (contenido 360).
		<i>tiled</i>	Valor binario para indicar si el contenido está dividido en tiles (contenido 360).
		<i>uri</i>	Identificador uniforme del origen del contenido multimedia o mulsemedia.
WEB	IDES, IDMS, Mulsemedia		Especifica la información necesaria para cualquier contenido relacionado disponible a través de la web.
		<i>id</i>	Valor de identificación único para el elemento.
		<i>protocol</i>	Protocolo utilizado (por ejemplo, http).
		<i>media_type</i>	Tipo de contenido (por ejemplo, website).
		<i>media_format</i>	Información del formato o codificación (por ejemplo, html5).
		<i>uri</i>	Identificador uniforme de recurso web.
LASTUPDATE	IDES, IDMS, Mulsemedia		Especifica la fecha de la última vez que el fichero ha sido actualizado.
		<i>media_format</i>	Formato de la fecha y hora.
		<i>value</i>	Valor de la fecha y hora.
CLOCK	IDES, IDMS, Mulsemedia		Especifica los metadatos para la tecnología de reloj global.
		<i>id</i>	Valor de identificación único para el elemento.
		<i>protocol</i>	Protocolo utilizado para la sincronización de relojes (por ejemplo, NTP).
		<i>media_type</i>	Tipo de contenido (por ejemplo, time).
		<i>media_format</i>	Información del formato o codificación (por ejemplo, YYYYMMDD-hhmmss).
		<i>metadata</i>	Breve descripción.
		<i>uri</i>	Identificador uniforme del servidor de reloj global.
IDMS	IDMS		Especifica la información necesaria en caso de existir un Manager IDMS.
		<i>id</i>	Valor de identificación único para el elemento.
		<i>protocol</i>	Protocolo utilizado para establecer la comunicación con el IDMS Manager (por ejemplo, WS).
		<i>metadata</i>	Breve descripción.
		<i>uri</i>	Identificador uniforme del IDMS Manager.

En la Fig. 17 se muestra un ejemplo de un fichero RMCF, en el que se señala la existencia de tres tipos diferentes de contenido accesible vía redes broadband (incluidos dentro de cada etiqueta *MEDIA*): contenido 360 u omnidireccional (*media type*= "360AV"), contenido audiovisual tradicional (*media type*= "AV") y contenido mulsemedia (*media type*= "SEM"), además del resto de etiquetas definidas en la tabla anterior.

```

<?xml version="1.0" encoding="UTF-8"?>
<Related Media Contents File>

  <MEDIA id="media0" media_type="360AV" media_format="h264/aac" metadata="360_cam"
  temi_init="3763721899965118886">
    <source protocol="http/dash" projection="ERP" tiled="false"
    uri="http://server.com/360/erp/stream.mpd"/>
  </MEDIA>

  <MEDIA id="media1" media_type="AV" media_format="h264/aac" metadata="TV_cam/front view"
  temi_init="3763721899965118886">
    <source protocol="http/dash" uri="http://server.com/front/stream.mpd"/>
  </MEDIA>

  <MEDIA id="mulsemedia0" media_type="SEM" media_format="xml" metadata="mulsemedia"
  temi_init="3763721899965118886">
    <source protocol="http" uri="http://server.com/mulsemedia/mmedia_effects.xml"/>
  <MEDIA/>

  <WEB id = "web1" protocol = "http" media_type = "website" media_format = "html5" metadata =
  "url_event/english" uri = "http://official_website.com/hybrid_IDMS_app"/>

  <IDMS id = "idms-mgr" protocol = "websocket" metadata = "idms manager uri" uri =
  "http://official_website.com:port/idms_manager1"/>

  <CLOCK id = "global-clock" protocol = "ntp" media_type = "time" media_format =
  "64_bit_ntp_time" metadata = "" uri = "ntp.specific_ntp_server.com"/>

  <LASTUPDATE protocol = "http" media_type = "time" metadata = "" format = "mm/dd/yyyy-
  hh:mm:ss" value = "31/12/1999-23:59:59"/>

</Related Media Contents File >

```

Fig. 17 Ejemplo de fichero RMCF

Al estar la señalización de contenidos complementarios incluida en el contenido principal, el dispositivo encargado de extraerla y analizarla será el dispositivo que ejecute la MS app, como por ejemplo la Smart TV). Dicho dispositivo debe ser capaz de analizar los descriptores TEMI del flujo MPEG2-TS, descargarse e interpretar el fichero RMCF. Tras conocer qué tipo de contenidos están disponibles durante la experiencia de consumo, la MS app enviará a las aplicaciones CS (CS apps) ejecutándose en el propio dispositivo (terminal híbrido) o en dispositivos secundarios con las que haya establecido previamente un canal de comunicación (canal *App2App*) los contenidos que sean compatibles y reproducibles por cada tipo de dispositivo conectado, así como información necesaria para la generación de los efectos multisensoriales si los hubiera. Como ejemplo, se podría estar enviando a una Tablet ejecutando la CS app, la URL del contenido generado por una cámara enfocando desde un ángulo diferente al de la cámara que generó el contenido principal (recibido por broadcast y siendo reproducido por la MS app), o enviando a un dispositivo generador o difusor de aromas los instantes de tiempo en los que se debe generar un aroma determinado.

4.4 TÉCNICAS DE AJUSTE DEL PROCESO DE REPRODUCCIÓN

En esta Tesis Doctoral, en todas las soluciones de sincronización propuestas, se han adoptado dos técnicas diferentes para ajustar los procesos de reproducción de las aplicaciones MS y CS a un punto de referencia determinado, con el fin de alcanzar un nivel de sincronización satisfactorio entre los mismos, ya sea en escenarios mono- (terminal híbrido) o multi- dispositivo, bien con dispositivos locales o remotos (entorno multidestinatarario). En particular, las dos técnicas son:

- Saltos y pausas (S&P): se trata de una técnica de ajuste abrupta, pero instantánea, puesto que el proceso de reproducción salta o se pausa una cierta cantidad de tiempo (según vaya atrasado o adelantado respecto a la referencia, respectivamente) hasta alcanzar a dicha referencia y así obtener el nivel de sincronización deseado. Esta técnica puede ser más perceptible por el consumidor y se adopta en esta Tesis Doctoral para realizar el primer ajuste iniciar la reproducción de un contenido determinado y cuando el nivel de asincronía entre contenidos es superior a un umbral definido (explicado en este capítulo más adelante).
- Ajuste adaptativo de la reproducción (AMP): se trata de una técnica de ajuste más suave, que consiste en acelerar o decelerar el proceso de reproducción (por ejemplo, presentando el contenido de vídeo a una tasa mayor o menor de fps con respecto a la tasa nominal del vídeo) dependiendo de si el proceso de reproducción va atrasado o adelantado respecto a la referencia, respectivamente). Esta técnica es menos perceptible por el consumidor y se adopta en esta Tesis Doctoral para llevar a cabo los ajustes necesarios durante la presentación del contenido, una vez ya se ha realizado el ajuste inicial basado en S&P. En el marco de investigación de esta Tesis Doctoral, la duración de los ajustes de tipo AMP que se aplican en el proceso CS, se calculan siguiendo la Ecuación 1:

Ecuación 1 Cálculo de la duración del ajuste adaptativo del proceso de reproducción AMP

$$t_{ajusteAMP} = abs\left(\frac{Valor\ asincronia}{Tasa\ reproducción\ CS}\right) \quad (1)$$

donde la tasa de reproducción CS deberá ser un valor distinto de 1 (valor de la tasa nominal) bien sea para acelerar (tasa > 1) o disminuir (tasa < 1) el proceso de reproducción.

4.5 DISEÑO Y ESPECIFICACIÓN DE UNA SOLUCIÓN DE SINCRONIZACIÓN DE LA REPRODUCCIÓN DE CONTENIDOS HÍBRIDOS EN UNO O VARIOS DISPOSITIVOS LOCALES

En la presente Tesis Doctoral se ha diseñado una serie de mecanismos y funcionalidades que permiten la sincronización de la reproducción de contenidos híbridos en uno o múltiples dispositivos, adoptando para tal fin un esquema M/S de control de la sincronización. En este esquema, la MS app actúa con el rol de máster y el resto de las CS apps, con el rol de esclavo. Cabe destacar que, en este escenario, la CS app puede estar ejecutándose en el mismo dispositivo que la MS app (terminal híbrido), o bien en otro dispositivo (en caso de tratarse de un escenario IDES), pues el mecanismo de sincronización es agnóstico en cuanto a la ubicación de la CS app (siempre y cuando las MS y CS apps estén ejecutándose en dispositivos conectados a la misma red), y, por tanto, es igualmente válido y se aplica del mismo modo tanto para escenarios mono- como multi- dispositivo.

Para alcanzar un estado de sincronización satisfactoria en este escenario, se deberá realizar el siguiente procedimiento por parte de la MS app: i) extracción del descriptor TEMI *timeline* del contenido recibido a través de la red broadcast; ii) monitorización de las tramas de vídeo hasta el instante previo a la renderización; iii) estimación del instante de presentación de la trama siendo renderizada. Se utiliza el instante de presentación estimado en lugar del instante de recepción,

para obtener una mayor precisión en el nivel de sincronización alcanzado, ya que al estimar dicho instante se tiene en cuenta la latencia que pueda tener el reproductor de contenido, desde que se recibe una trama de vídeo hasta el instante de presentación al usuario. Dicho instante estimado será procesado y calculado en un formato basado en NTP, el cual, junto con los valores de tiempo almacenados en el descriptor TEMI *timeline*, que también tiene un formato basado en NTP, serán utilizados por la MS app (máster) para obtener la referencia temporal que proporcionará al resto de CS apps (esclavos). La transmisión de esta información temporal se realiza a través del canal de comunicación entre procesos o MS y CS apps, denominado *App2App* en el estándar HbbTV [5] y cuyo proceso de creación y establecimiento está definido en DIAL [30], tal y como se ha explicado en el Capítulo 2. El envío de dicha información se realizará de forma periódica con una frecuencia configurable (por ejemplo, cada 2 segundos) según las características del caso de uso o escenario particular en el que se necesite la solución IDES. La Fig. 18 muestra el diagrama de flujo del proceso seguido en el dispositivo principal encargado de ejecutar la MS app. Una vez se inicia el proceso, se establece la conexión con el servidor DIAL y se sincroniza el reloj del proceso con el reloj global. Una vez la MS app recibe contenido broadcast, inicia su reproducción. Tras esto y de forma paralela, por un lado, extrae y analiza los descriptors TEMI insertados en la trama. Esta extracción y análisis se lleva a cabo cada un número determinado de tramas, coincidiendo con el tipo de trama I. Tras la extracción y análisis de los descriptors TEMI, envía la información temporal a las CS app que ya estén conectadas. Por otro lado, la MS app está a la espera escuchando posibles peticiones de nuevas conexiones de CS apps. En el caso de haber una nueva conexión, la MS app envía información respecto los contenidos disponibles para la nueva CS app y pasa a enviarle, periódicamente, la información temporal necesaria.

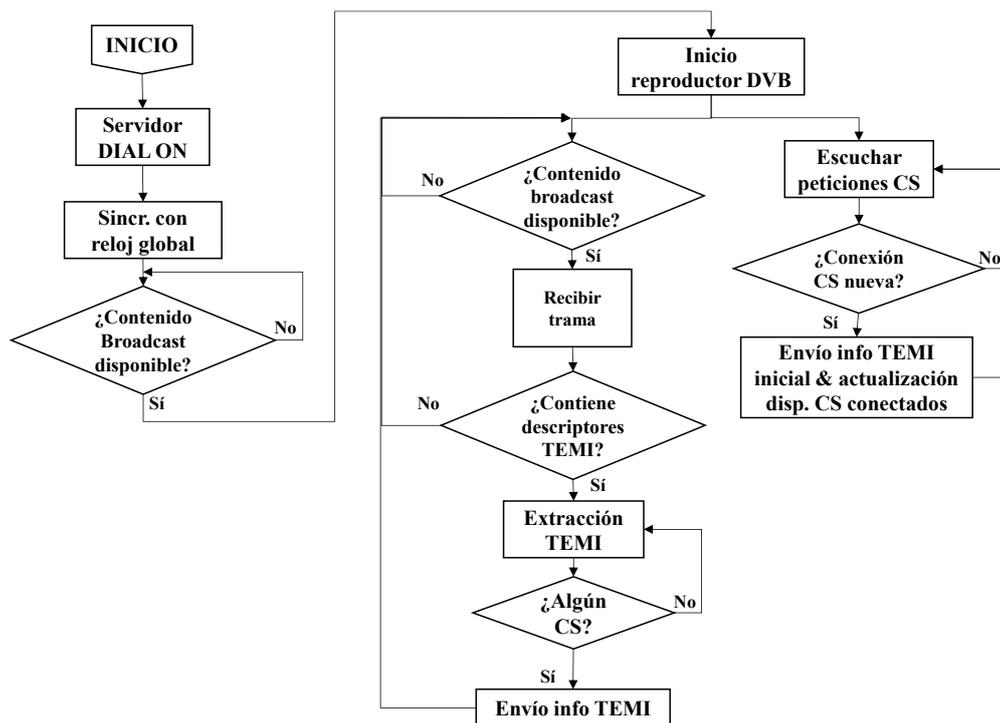


Fig. 18 Diagrama de flujo del dispositivo principal ejecutando la MS app

Cabe destacar que en la Fig. 18 existe un contador de tramas que permite la lectura del descriptor TEMI *timeline* de manera periódica, con el fin de no sobrecargar el proceso con una lectura continuada de esta información.

En la Fig. 19 se muestran los campos existentes en el mensaje de sincronización transmitido por la MS app a las CS apps. En particular, este mensaje consta de dos campos, en los cuales se

almacena el instante de generación de la trama de contenido que se está reproduciendo junto con el instante de reproducción actual (según la referencia de tiempo de global).

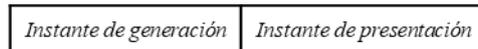


Fig. 19 Campos del mensaje de sincronización IDES

En los tipos de escenarios en los que se centra la Tesis Doctoral, al tomar la decisión de adoptar mecanismos y protocolos especificados en el estándar HbbTV, se está utilizando el canal de comunicación *App2App* que está basado en la tecnología *Websocket*, que, a su vez, se basa en comunicaciones unicast. Este proceso de envío de mensajes de la MS app a las CS apps podría optimizarse si se utilizara UDP/multicast, pero se descartó por no ser una tecnología propuesta en el estándar HbbTV.

Una vez la CS app ha recibido el mensaje de sincronización IDES con la información temporal del estado de reproducción de la MS app, es capaz de analizar dicha información y compararla con el estado de su propio proceso reproductor del contenido complementario. La CS app calcula el nivel de asincronía respecto al proceso reproductor de la MS app y, en caso de ser necesario (cuando dicho nivel supera un umbral predefinido), deberá realizar un ajuste suave (AMP) en su proceso de reproducción.

En la Fig. 20, se muestra la relación entre las marcas temporales de las MS y CS apps utilizadas en el mecanismo de sincronización propuesto para los escenarios mono- y multidispositivo (cuando estos están en local). Los instantes 'm' y 'n' se corresponden, respectivamente, con los instantes de presentación y generación correspondientes a la trama de contenido que se estaba reproduciendo en la MS app cuando esta envió el mensaje de sincronización IDES a la CS app. Este par de marcas indican que en el instante 'm', se estaba visualizando la trama del contenido generada en el instante 'n'. De forma análoga, los instantes 'u' y 'v', se refieren a la trama que se estaba reproduciendo por el proceso de reproducción de la CS app en el momento de recibir dicho mensaje.

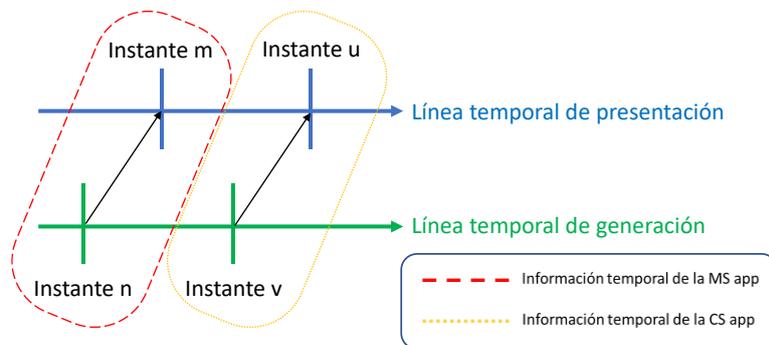


Fig. 20 Relación entre las marcas temporales utilizadas en el mecanismo de sincronización propuesto para escenarios mono- y multidispositivo (locales)

La CS app, para calcular el nivel de sincronización (si existe asincronía) de su proceso de reproducción con respecto al de la MS app, utiliza las Ecuaciones 2 y 3:

Ecuación 2 Cálculo del instante de generación objetivo

$$\begin{aligned} \text{Instante de generación objetivo} \\ = \text{instante } v + (\text{instante } m - \text{instante } n) \end{aligned} \quad (2)$$

Ecuación 3 Cálculo del nivel de sincronización alcanzado por una CS app en un escenario mono- o multidispositivo

$$\text{Asincronía} = \text{Instante de generación objetivo} - \text{instante } u \quad (3)$$

Inicialmente se calcula el valor del instante de generación de la trama que debería estar reproduciendo la CS app en caso de que su proceso de reproducción estuviera completamente sincronizado con el de la MS app (*Instante de generación objetivo*). Para ello utiliza tanto información de su propio proceso como la información temporal del proceso de la MS app obtenida en el mensaje de sincronización IDES. Dicho valor se utiliza para calcular la asincronía entre ambos procesos de reproducción (diferencia entre el *Instante de generación objetivo* y el instante de generación de la trama reproduciéndose en la CS app). Si el valor de *Asincronía* es negativo, significa que el proceso de reproducción de la CS app se encuentra atrasado respecto al proceso de reproducción de la MS app. Por el contrario, si el resultado es positivo, el proceso de reproducción de la CS app se encuentra adelantado respecto al del proceso de reproducción de la MS app.

Para evitar que el proceso de reproducción de la CS app entre en una situación no deseada de ajustes continuos, se propone el uso de un temporizador entre ajustes consecutivos. No se permitirán nuevos ajustes hasta pasado un tiempo después de realizar un ajuste. Así se permite que el proceso de reproducción pueda llegar a un estado estable tras haber realizado un ajuste antes de volver a comparar con el proceso de reproducción de la MS app. En la Fig. 21 se muestra el diagrama de flujo del comportamiento de un dispositivo que ejecuta la CS app (ya sea el propio terminal híbrido o un dispositivo secundario).

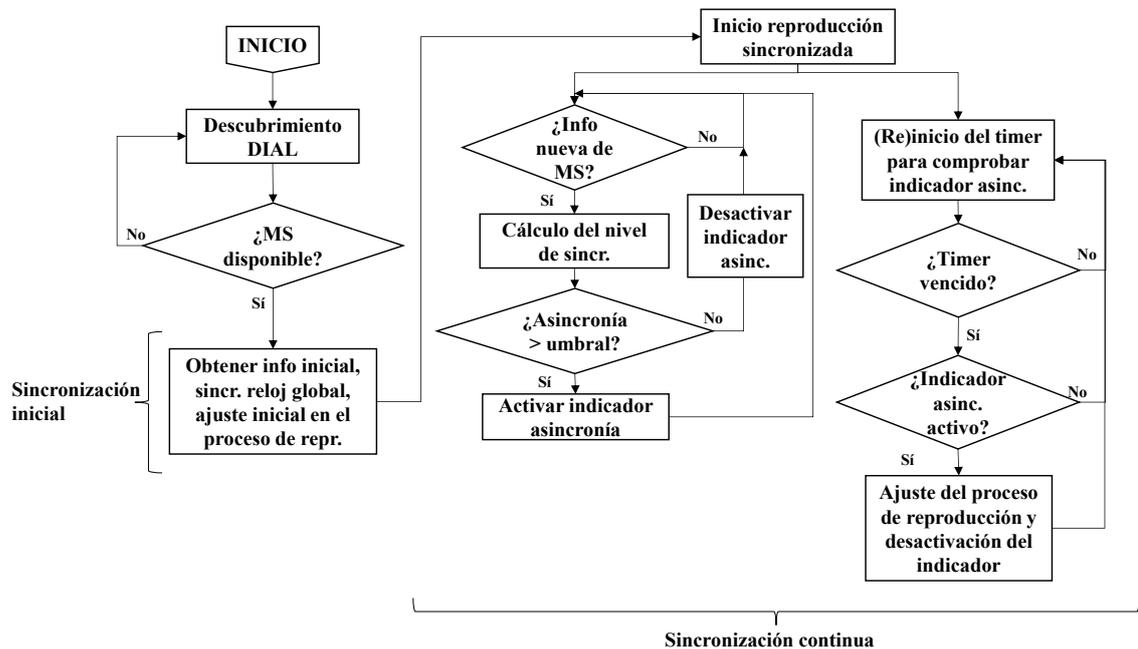


Fig. 21 Diagrama de flujo de dispositivo ejecutando la Companion Screen app

Tal y como se puede observar en el diagrama de flujo de la Fig. 21, el mecanismo de sincronización se puede dividir en dos fases (similar a otros trabajos como [76], [56] o [59]): una fase de sincronización inicial y una fase posterior de sincronización continua (más precisa). La *fase de sincronización inicial*, tal y como su propio nombre indica, se lleva a cabo al iniciar la reproducción en la CS app, momento en el cual la MS app se supone que ya estará reproduciendo contenidos. Esta fase consiste en realizar un salto inicial en la reproducción de la CS app hasta un instante específico del contenido, para alcanzar un estado de reproducción inicial sincronizado

con el estado del proceso de reproducción de la MS app. A partir de dicho momento, se inicia la *fase de sincronización continua*, que se llevará a cabo durante toda la reproducción del contenido. Esta fase consiste en la monitorización constante (con una cierta frecuencia configurable) del nivel de asincronía entre el proceso de reproducción de la CS app respecto al proceso de reproducción de la MS app. En el caso en el que se exceda un determinado umbral de la asincronía máxima tolerable, la CS app deberá realizar un ajuste en su proceso de reproducción siempre y cuando se encuentre dentro de la ventana de tiempo definida por un temporizador (timer en la parte derecha del diagrama en la Fig. 21) que evita que dos ajustes consecutivos se solapen en el tiempo.

4.6 DISEÑO Y ESPECIFICACIÓN DE UNA SOLUCIÓN DE SINCRONIZACIÓN DE LA REPRODUCCIÓN DE CONTENIDOS HÍBRIDOS EN VARIOS DISPOSITIVOS REMOTOS

La tecnología de transmisión broadcast es unidireccional, por lo que en experiencias de consumo de contenido de TV compartidas entre consumidores remotos (es decir, entre dispositivos remotos) que requieran de una solución de sincronización IDMS, serán necesarios mecanismos para la creación y gestión de grupos de usuarios (o dispositivos) remotos, así como un canal de comunicaciones bidireccional entre los dispositivos remotos, fácil de implementar a través de redes IP (broadband).

En este apartado primero se definen una serie de mensajes de control con el fin de crear, unirse o abandonar una sesión de consumo compartida por varios dispositivos remotos, así como también para alcanzar un estado de sincronización entre destinatarios (IDMS) satisfactorio. En segundo lugar, se presenta el diseño y la especificación de dos soluciones IDMS para alcanzar un nivel de sincronización satisfactorio entre los procesos de reproducción de los dispositivos de los destinatarios remotos. Una se basa en el uso del esquema de control de sincronización centralizado SMS y la otra en el esquema de control de sincronización distribuido DCS, ambos explicados a continuación. El uso de un esquema de control de sincronización centralizado M/S ha sido descartado debido a que dicho esquema para escenarios IDMS requiere de mecanismos adicionales para conseguir un nivel de sincronización satisfactorio, como por ejemplo, una monitorización minuciosa para evitar vaciados o desbordamientos del búfer de reproducción cuando el dispositivo con rol de máster presenta un instante del contenido muy por delante o muy por detrás de otros clientes con rol de esclavo, respectivamente [77].

La Fig. 22 muestra los dos esquemas de control de sincronización SMS y DCS para los que se han diseñado mecanismos de sincronización IDMS. En ambos esquemas se hace uso de un mensaje denominado *IDMS Report*, que contiene la información temporal de cada uno de los clientes involucrados en la sesión. Adicionalmente, en el esquema de control SMS, se define un tipo de mensaje adicional, denominado *IDMS Settings*, que contiene información temporal que servirá como de referencia a la cual se deben de ajustar los procesos de reproducción de los dispositivos cliente involucrados en la sesión. Ambos tipos de mensaje se definen más exhaustivamente en las siguientes secciones de este Capítulo. Además, existe una entidad adicional en el esquema de control SMS, tal y como se puede observar en la Fig. 22, denominada *IDMS Manager*.

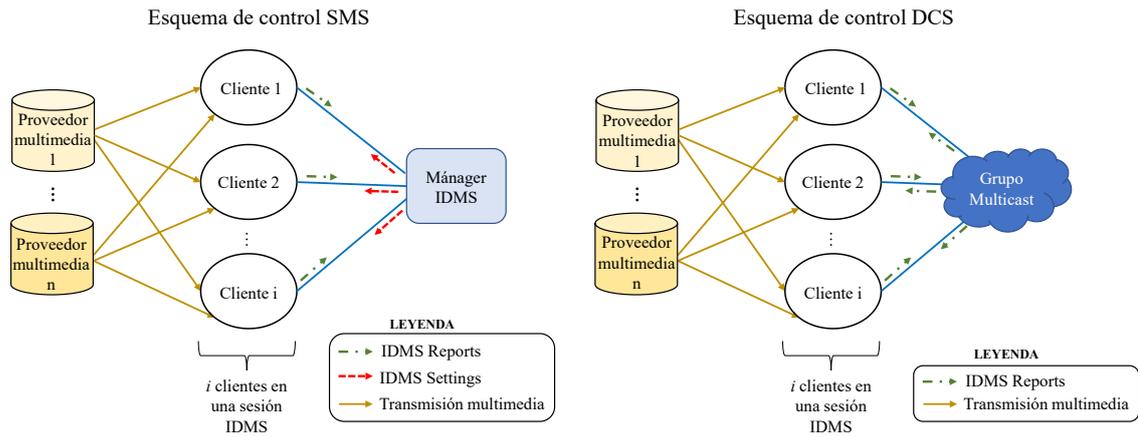


Fig. 22 Esquemas de control de sincronización en el escenario inter-destinatario (centralizado en la izquierda y distribuido en la derecha)

PARTICULARIDADES EXISTENTES EN EL ESQUEMA DE CONTROL SYNCHRONIZATION MAESTRO SCHEME

En un esquema de control SMS el IDMS Manager, cuyos módulos se muestran en la Fig. 23, es la entidad encargada de recibir la información de los Clientes IDMS conectados en una sesión, y de calcular y proporcionar a todos ellos una referencia temporal para la sincronización IDMS. Además, tiene la capacidad de gestionar sesiones independientes de manera simultánea. Conoce el estado de cada Cliente IDMS en cada sesión, y es capaz de enviar un mensaje que implique un ajuste en el proceso de reproducción en los clientes IDMS en cada sesión IDMS activa, seleccionando además una estrategia determinada de selección de una referencia de sincronización. Todo esto se realiza a través de una serie de intercambios de mensajes del tipo *IDMS Reports* e *IDMS Settings*, descritos en este capítulo.

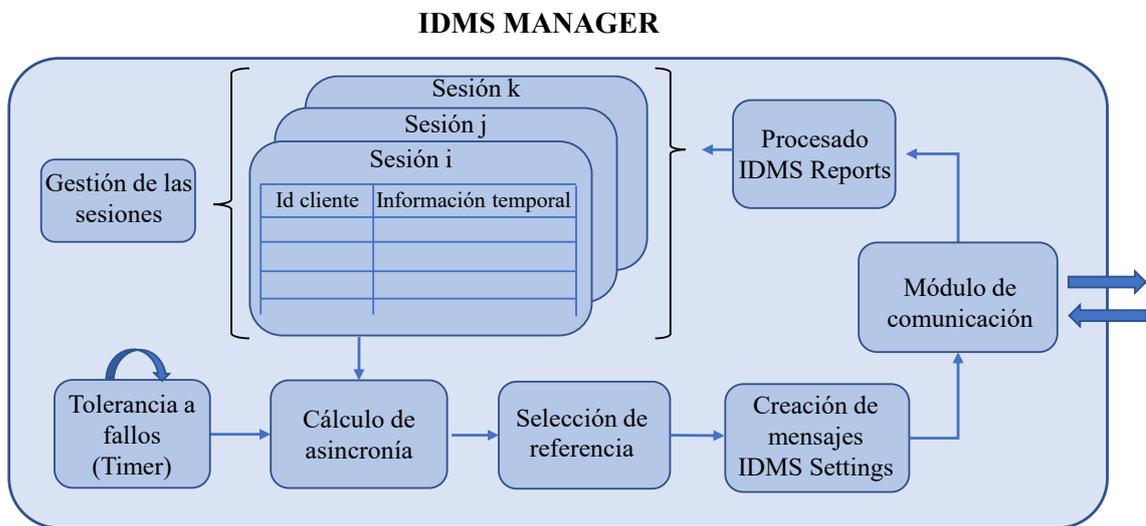


Fig. 23 Módulos integrados dentro del IDMS Manager

A continuación, se definen los mensajes de control y los procedimientos diseñados necesarios para que un Cliente IDMS pueda crear una sesión, unirse a una sesión ya existente o abandonar una sesión. Además, también se definen los mensajes involucrados en el proceso de sincronización IDMS.

4.6.1 MENSAJES DE SINCRONIZACIÓN INTER-DESTINATARIO PARA EL CONTROL DE LA SESIÓN Y DE LA SINCRONIZACIÓN

Los mensajes de control (de la sesión y de la sincronización) consisten en mensajes de tipo textual (cadena de caracteres) y han sido diseñados para ser compatibles con cualquier tipo de esquema de control de la sincronización, aunque con alguna ligera variación en algunos campos específicos para cada esquema.

La sintaxis definida para el conjunto de mensajes sigue la estructura presentada en la Fig. 24. Los campos son cadenas de texto separadas por el carácter ';'. Los tres primeros campos son comunes para todos los tipos de mensajes. El primer campo identifica el tipo de mensaje (valor numérico); el segundo contiene un identificador asociado al Cliente IDMS que envía el mensaje (cadena de caracteres); y finalmente, el tercero almacena un valor checksum de comprobación (siguiendo el procedimiento definido en la RFC 1071 [149]), calculado a partir de la información de los dos primeros campos.

<i>Tipo mensaje IDMS</i>	<i>Id cliente</i>	<i>Checksum</i>	<i>Otros campos (si necesarios)</i>
--------------------------	-------------------	-----------------	-------------------------------------

Fig. 24 Formato genérico de los mensajes IDMS propuestos

MENSAJES DE SINCRONIZACIÓN INTER-DESTINATARIO PARA EL CONTROL DE LA SESIÓN

En la Tabla 3 se resumen los mensajes de control de la sesión IDMS, así como el tipo de esquema de control para el que se aplica (SMS o DCS) y la dirección de la transmisión. Cada uno de ellos se explica a continuación.

Tabla 3 Conjunto de mensajes para el control de sesiones definidos para escenarios IDMS

Mensaje IDMS	Esquema	Dirección	Propósito
Petición Hello	SMS & DCS	IDMS Manager → Clientes IDMS (SMS) Cliente IDMS → Clientes IDMS (DCS)	Comprobar conectividad. En el esquema DCS también se utiliza para crear una sesión.
Respuesta Hello	SMS & DCS	Cliente IDMS → IDMS Manager (SMS) Cliente IDMS → Clientes IDMS (DCS)	Respuesta a petición Hello.
Crear sesión	SMS	Cliente IDMS → IDMS Manager	Crear una nueva sesión
ACK session creada	SMS	IDMS Manager → Cliente IDMS	Confirmación de session creada
Join	SMS & DCS	Cliente IDMS → IDMS Manager (SMS) Cliente IDMS → Clientes IDMS (DCS)	Informar al resto de participantes de una session que un nuevo cliente se une.
Leave	SMS & DCS	Cliente IDMS → IDMS Manager (SMS) Cliente IDMS → Clientes IDMS (DCS)	Informar al resto de clientes que se abandona la sesión.

- *IDMS Hello Request*

Este tipo de mensaje (junto al siguiente en la lista) se utiliza para comprobar la conectividad. En particular lo envía el IDMS Manager a los Clientes IDMS para verificar que están conectados si se adopta un esquema SMS y se utiliza para crear una nueva sesión si el esquema adoptado es el DCS (explicado en 4.6.2). Contiene únicamente los campos comunes presentados en la Fig. 24. El campo tipo de mensaje contiene el valor 1.

- *IDMS Hello Response*

Cuando un Cliente IDMS recibe un mensaje del tipo *Hello Request*, debe responder, de manera inmediata, con un mensaje *Hello Response* (para el caso con esquema de control DCS, únicamente si no se ha recibido antes una respuesta de otro Cliente IDMS). Este mensaje contiene sólo los campos de la Fig. 24. El campo tipo de mensaje contiene el valor 2.

El mensaje *IDMS Hello Response* se envía de manera repetida (con un número de repeticiones configurable) en un periodo de tiempo corto, con el fin de que alguna de las copias llegue al Cliente IDMS transmisor del mensaje *IDMS Hello Request*, por si ha habido algún tipo de pérdida de paquetes.

- *IDMS Create Session*

Un Cliente IDMS envía este mensaje al IDMS Manager para crear una sesión o grupo IDMS. Contiene los campos comunes de la Fig. 24, con el valor 3 en el campo tipo de mensaje.

- *IDMS Create Session ACK*

Este mensaje lo envía el IDMS Manager como respuesta al anterior y contiene los tres campos comunes, además de un cuarto campo que contiene el identificador de la sesión o grupo (en adelante, *Group id*), Fig. 25. El campo tipo de mensaje contiene el valor 4. En caso de que la sesión no haya podido crearse correctamente (por ejemplo, que el IDMS Manager haya sufrido un error al configurar la sesión), el campo *ID Grupo* contendrá un valor de tipo *null*.

<i>IDMS Create Session ACK</i>	<i>Id cliente</i>	<i>Checksum</i>	<i>ID Grupo</i>
--------------------------------	-------------------	-----------------	-----------------

Fig. 25 Formato del mensaje IDMS Create Session ACK

- *IDMS Join*

Este mensaje (Fig. 26) lo envía un cliente IDMS y contiene los tres campos comunes (en este caso, el campo tipo de mensaje contiene el valor 5), además de: i) un campo donde se almacena el identificador del grupo al que el Cliente IDMS que envía este mensaje quiere unirse (*ID Grupo*); ii) un campo donde indica el número total de ítems o campos adicionales incluidos en este mensaje; y iii) una lista de ítems que aportan información adicional sobre el Cliente IDMS que quiere unirse a la sesión. En particular, los ítems están basados en el mensaje *RTCP Source Description* (SDES), definidos en [2] y tienen el formato indicado en la Fig. 27. La Tabla 4 presenta los ítems definidos, de entre los cuales destaca el denominado como NAME, que es el único que debe incluirse obligatoriamente.

<i>IDMS Join</i>	<i>Id cliente</i>	<i>Checksum</i>	<i>ID Grupo</i>	<i>Núm. de ítems</i>	<i>Listado de ítems [...]</i>
------------------	-------------------	-----------------	-----------------	----------------------	-------------------------------

Fig. 26 Formato del mensaje IDMS Join

<i>Id ítem</i>	<i>longitud</i>	<i>payload</i>
----------------	-----------------	----------------

Fig. 27 Formato de los ítems incluidos dentro del mensaje IDMS Join

Tabla 4 Ítems disponibles para el mensaje IDMS Join

ítem	id	Descripción
NAME	1	[Obligatoria] Nombre del usuario
EMAIL	2	[Opcional] Email del usuario
PHONE	3	[Opcional] Número telefónico del usuario
LOC	4	[Opcional] Localización geográfica del usuario
TOOL	5	[Opcional] Nombre de la aplicación

- *IDMS Leave*

Este mensaje (Fig. 28) contiene los tres campos comunes (en este caso, el campo tipo de mensaje contiene el valor 6), además de un cuarto campo donde se identifica la sesión que se quiere abandonar.

<i>IDMS Leave</i>	<i>Id cliente</i>	<i>Checksum</i>	<i>ID Grupo</i>
-------------------	-------------------	-----------------	-----------------

Fig. 28 Formato del mensaje IDMS Leave

MENSAJES DE SINCRONIZACIÓN INTER-DESTINATARIO PARA EL CONTROL DE LA SINCRONIZACIÓN

En la Tabla 5 se resumen los mensajes de control de la sincronización IDMS, así como el tipo de esquema de control para el que se aplica (SMS o DCS) y la dirección de la transmisión. A continuación, se describe cada uno de ellos.

Tabla 5 Conjunto de mensajes para el control de sincronización definidos para escenarios IDMS

Mensaje IDMS	Esquema de control	Dirección	Propósito
Report	SMS & DCS	Cliente IDMS → IDMS Manager (SMS) Cliente IDMS → Clientes IDMS (DCS)	Envío de información temporal para el cálculo del valor de sincronización en la sesión.
Settings	SMS	IDMS Manager → Clientes IDMS	Informar a los Clientes IDMS que es necesario un ajuste en el proceso de reproducción.

- *IDMS Report*

Este mensaje contiene los tres campos comunes a todos los mensajes (en este caso, el campo tipo de mensaje contiene el valor 7), además de cinco campos más (Fig. 29): el cuarto campo incluye el id de la sesión (*id sesión*); el quinto campo almacena el valor de la ronda de sincronización actual (explicado más adelante); y los tres últimos campos contienen marcas temporales (con formato NTP) indicando la marca temporal o timestamp de generación insertado en el contenido, el instante de reproducción o presentación de dicho contenido y el instante de transmisión del mensaje.

<i>IDMS Report</i>	<i>Id cliente</i>	<i>Checksum</i>	<i>Campos adicionales de IDMS Report</i>
--------------------	-------------------	-----------------	--

Campos adicionales de IDMS Report:

<i>Id sesión</i>	<i>n° R_{seq}</i>	<i>Marca temporal de contenido</i>	<i>Marca temporal de presentación</i>	<i>Marca temporal de transmisión</i>
------------------	---------------------------	------------------------------------	---------------------------------------	--------------------------------------

Fig. 29 Formato del mensaje IDMS Report

- *IDMS Settings*

Este mensaje es similar al *IDMS Report*, salvo por dos diferencias (Fig. 30): en este caso, el campo tipo de mensaje contiene el valor 8 y en este mensaje no existe un campo para almacenar el instante en el que se envía el mensaje.

<i>IDMS Settings</i>	<i>Id cliente</i>	<i>Checksum</i>	<i>Campos adicionales de IDMS Settings</i>
----------------------	-------------------	-----------------	--

Campos adicionales de IDMS Settings:

<i>Id sesión</i>	<i>n° R_{seq}</i>	<i>Marca temporal de contenido</i>	<i>Marca temporal de presentación</i>
------------------	---------------------------	------------------------------------	---------------------------------------

Fig. 30 Formato del mensaje IDMS Settings

4.6.2 CREACIÓN DE UNA SESIÓN DE SINCRONIZACIÓN INTER-DESTINATARIO

El procedimiento para crear una sesión IDMS dependerá del esquema de control de sincronización seleccionado. Por un lado, si se utiliza el esquema SMS, el mensaje *IDMS Create Session* es transmitido por el Cliente IDMS que quiere crear la sesión al IDMS Manager. Cabe destacar que la ubicación del IDMS Manager es conocida por el Cliente IDMS al estar debidamente señalizada en el fichero RMCF (ver el ejemplo de la Fig. 17). Al recibir este mensaje, el IDMS Manager le responde con un mensaje del tipo *IDMS Create Session ACK*, con el objetivo de notificar si la sesión se ha creado correctamente. El campo *id session* del mensaje *IDMS Create Session ACK* contendrá un valor único para la sesión creada o *null*, dependiendo de si se ha creado correctamente la sesión o no, respectivamente.

Por otro lado, si se utiliza un esquema de control DCS, se emplean comunicaciones multicast a una dirección IP de grupo (multicast) que conocen todos los miembros de la sesión. Dicha dirección IP multicast se utiliza como el identificador de la sesión IDMS. Cuando un Cliente IDMS quiere crear una nueva sesión, necesitará encontrar una IP multicast que no esté siendo ocupada por otra sesión en un rango de direcciones multicast disponibles (definir la gestión de este tipo de direcciones por parte del proveedor del servicio de TV híbrida con funcionalidades IDMS queda fuera del alcance de la presente Tesis Doctoral). Tras seleccionar una dirección IP multicast, el Cliente IDMS envía un *IDMS Hello Request* encapsulado en un datagrama cuya dirección de destino es dicha dirección IP multicast, con el fin de comprobar si la dirección multicast está siendo utilizada en otra sesión activa o no. Si ya está siendo utilizada en otra sesión, los Clientes IDMS de dicha sesión deberán responder enviando el mensaje *IDMS Hello Response*. Si se reciben estos mensajes, se deberá buscar otra dirección IP multicast y repetir el proceso hasta encontrar una que esté libre en la que no recibirá respuestas de otros clientes IDMS.

Una vez creada la sesión IDMS, el Cliente IDMS que la ha creado puede compartir (por ejemplo, a través de diferentes vías o aplicaciones como Whatsapp, Telegram, Twitter, e-mail, etc.) la información necesaria (el *Id sesión*, si el esquema de control es SMS, o la dirección IP multicast, en el caso del esquema de control DCS) para que otros Clientes IDMS puedan unirse a la sesión. Cuando un Cliente IDMS se une correctamente a una sesión, también puede compartir esta información con otros clientes para invitarles a unirse a dicha sesión.

La Fig. 31 muestra el diagrama de flujo de un Cliente IDMS que crea una sesión bajo un esquema de control DCS.

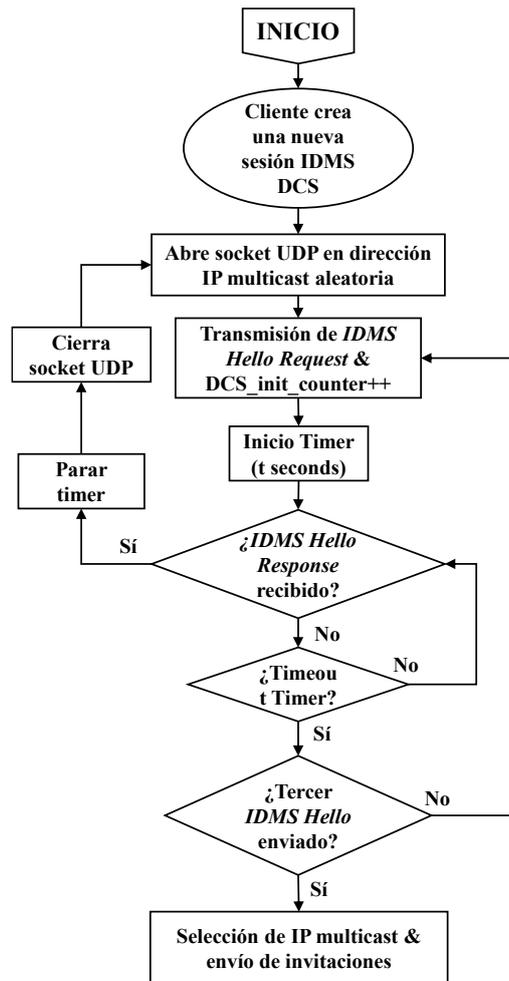


Fig. 31 Diagrama de flujo seguido por un cliente IDMS para crear una sesión IDMS bajo un esquema de control DCS

4.6.3 UNIRSE A UNA SESIÓN DE SINCRONIZACIÓN INTER-DESTINATARIO

Independientemente del esquema de control utilizado en una determinada sesión, un Cliente IDMS debe enviar un mensaje *IDMS Join* para poder unirse a dicha sesión. En el caso en el que el esquema de control sea SMS, el Cliente IDMS envía este mensaje al IDMS Manager, especificando el *id session* de la sesión a la que desea unirse, además de la información definida en el apartado 0. Una vez el IDMS Manager ha recibido esta información, responde al Cliente IDMS con un mensaje del tipo IDMS Settings. De esta manera, el Cliente IDMS asume que se ha podido unir a la sesión correctamente y, además, conoce el punto de reproducción al que debe sincronizarse para llegar a un estado de sincronización IDMS inicial satisfactorio cuanto antes.

En el caso en el que el esquema de control sea DCS, el Cliente IDMS envía el mensaje *IDMS Join* a la dirección IP multicast que se esté utilizando en dicha sesión, incluyendo la información pertinente. En este caso, el campo *id session* no se utiliza, pues no es necesario.

4.6.4 ABANDONAR UNA SESIÓN DE SINCRONIZACIÓN INTER-DESTINATARIO

Para informar que un Cliente IDMS desea abandonar una sesión se utilizará el mensaje *IDMS Leave*. En un esquema de control SMS, este mensaje se enviará al IDMS Manager. Una vez ha recibido el mensaje, el IDMS Manager identifica al Cliente IDMS que desea abandonar la sesión y descarta toda la información asociada al mismo que tenía almacenada y que estaba utilizando para obtener la referencia para la sincronización IDMS a enviar a los clientes IDMS y la vuelve a calcular sin tenerla en cuenta.

Respecto al esquema de control DCS, el Cliente IDMS que desea abandonar la sesión enviará el mensaje *IDMS Leave* a la dirección IP multicast utilizada en dicha sesión. Tras esto, el resto de Clientes IDMS pueden identificar al Cliente que abandona la sesión y, así, descartar toda la información asociada al mismo que tengan almacenada y que estaban utilizando para obtener la referencia para la sincronización IDMS.

4.6.5 SOLUCIÓN DE SINCRONIZACIÓN DE SINRONIZACIÓN INTER-DESTINATARIO

A continuación, se exponen las soluciones de sincronización IDMS basadas en un esquema de control centralizado (SMS) o distribuido (DCS).

SOLUCIÓN DE SINCRONIZACIÓN IDMS BASADA EN UN ESQUEMA DE CONTROL

SYNCHRONIZATION MAESTRO SCHEME

En el esquema SMS, el IDMS Manager es el encargado de calcular el nivel de asincronía existente entre los procesos de reproducción de los Clientes IDMS involucrados en una sesión, procesando la información temporal contenida en los mensajes *IDMS Report* recibidos de cada uno de ellos. Tan pronto como recibe el primer *IDMS Report*, el IDMS Manager inicializa un temporizador para controlar un período de tiempo dentro del cual debe admitir mensajes *IDMS Report* provenientes del resto de Clientes IDMS de la sesión. Dicho periodo de tiempo se denomina ronda de sincronización o *Sync Round*, y finaliza cuando finaliza el temporizador, o bien cuando el IDMS Manager ha recibido los mensajes *IDMS Report* de todos los Clientes IDMS de la sesión. En particular, se propone que el periodo *Sync Round* se calcule a partir de la Ecuación 4:

Ecuación 4 Cálculo de la ronda de sincronización

$$Sync\ Round = 2 * P_{IDMS} + Max(e2e_{delay_i}) \quad (4)$$

, donde P_{IDMS} es el periodo de tiempo (configurable) entre dos envíos consecutivos de mensajes *IDMS Report*, y $e2e_{delay_i}$ es el retardo extremo-a-extremo entre el i -ésimo Cliente IDMS y el IDMS Manager.

Tras unirse satisfactoriamente a una sesión IDMS, el Cliente IDMS, cada $P_{IDMS} \pm \Delta$ segundos, envía un mensaje *IDMS Report* al IDMS manager, donde Δ permite pseudo-aleatorizar el periodo de transmisión con el fin de evitar posibles sobrecargas o congestión en la red, puesto que el IDMS manager podría recibir todos los mensajes *IDMS Reports* de todos los clientes en el mismo instante.

Una vez ha finalizado el período *Sync Round* o se han recibido los mensajes *IDMS Report* de todos los clientes IDMS, el IDMS manager calcula la asincronía de la sesión (*session async*), correspondiente al valor de asincronía existente entre el punto de reproducción de referencia y el punto de reproducción del proceso de reproducción cuya asincronía sea mayor dentro de la misma sesión. Si la asincronía de la sesión (*session async*) excede un valor umbral de asincronía máxima aceptable (denominado *session threshold*), el IDMS Manager calculará el punto de reproducción de referencia (*sync reference*) y lo enviará a los Clientes IDMS en un mensaje *IDMS Settings*. La Fig. 32 describe el funcionamiento del IDMS Manager.

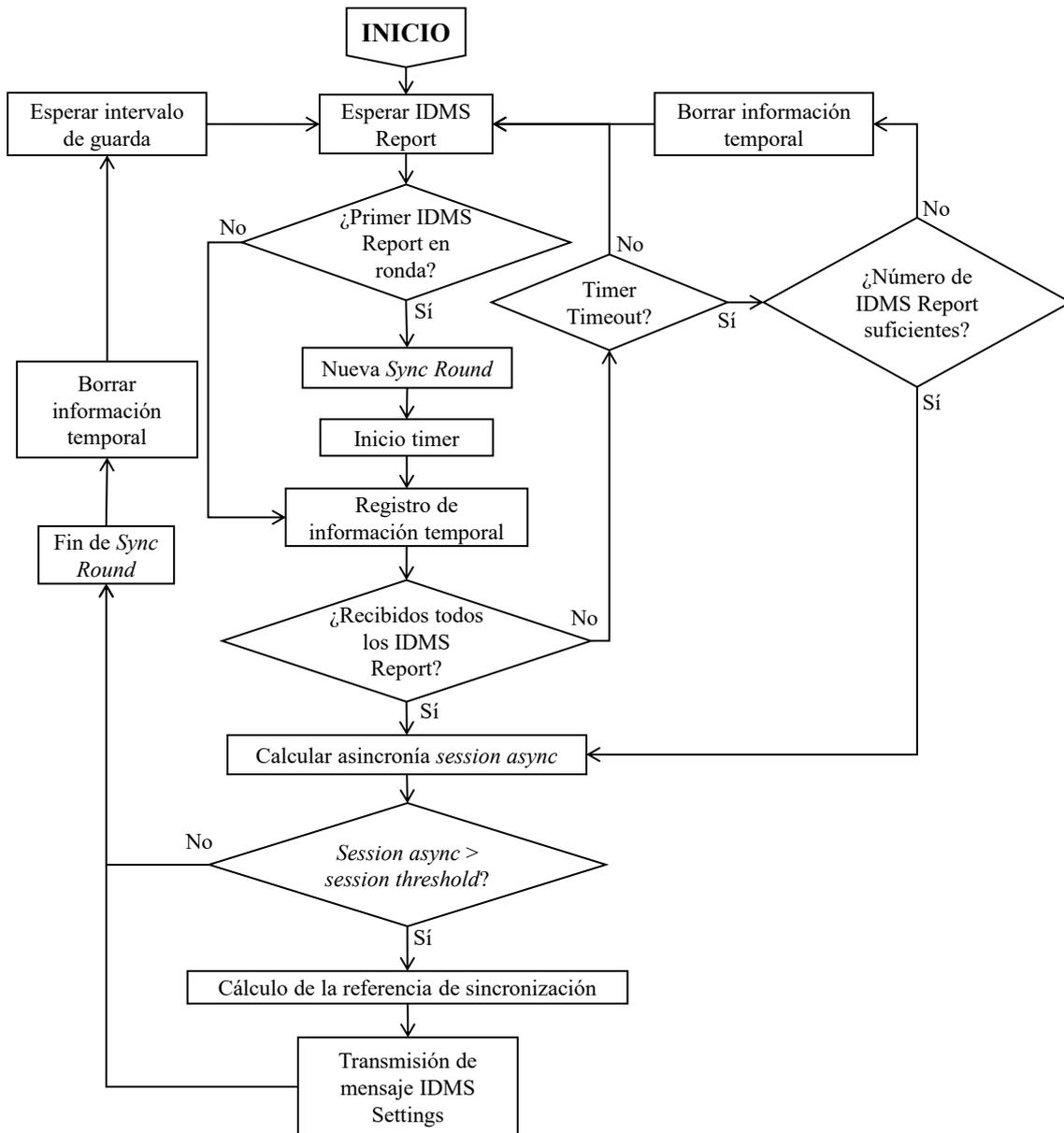


Fig. 32 Diagrama de flujo del proceso seguido en el IDMS Manager en el esquema de control SMS

Durante la reproducción del contenido, el Cliente IDMS extrae cada P_{IDMS} segundos las marcas temporales insertadas en el descriptor TEMI timeline del flujo broadcast, con el fin de generar y transmitir el mensaje IDMS Report. Cuando el Cliente recibe un mensaje *IDMS Settings*, calculará la asincronía de su propio proceso de reproducción respecto al punto de referencia que ha recibido en el mensaje *IDMS Settings*, denominada *IDMS internal async*. Si el valor de *IDMS internal async* calculado supera el umbral de asincronía máxima aceptable (*IDMS internal async threshold*), el Cliente IDMS deberá realizar un ajuste en su proceso de reproducción para alcanzar un estado de sincronización IDMS satisfactorio. La Fig. 33 muestra el diagrama de flujo del comportamiento de un Cliente IDMS bajo un esquema de control SMS. Cabe destacar que, con el propósito de simplificar el diagrama, se ha omitido el proceso de creación o unión a una sesión IDMS puesto que se trata de un proceso lineal de peticiones y respuestas Cliente – Servidor en las que no ocurre ninguna otra operación de forma paralela y este proceso ha quedado definido en el apartado 0.

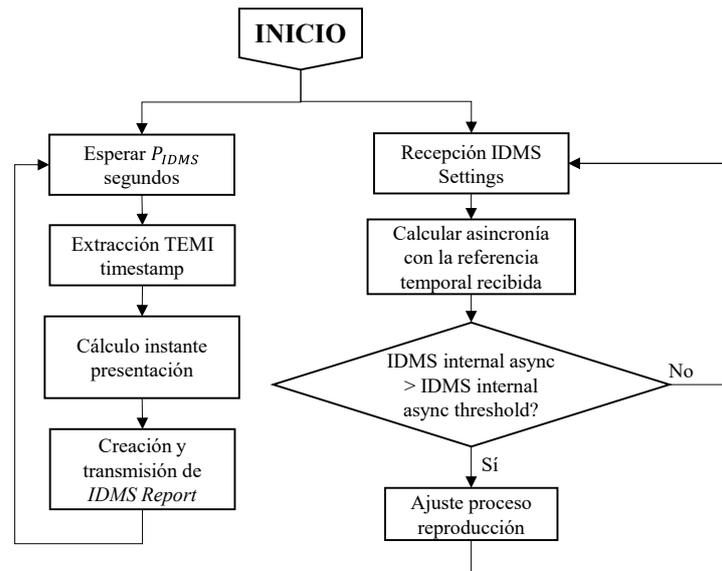


Fig. 33 Diagrama de flujo de un Cliente IDMS bajo un esquema de control SMS

SOLUCIÓN DE SINCRONIZACIÓN DE SINCRONIZACIÓN INTER-DESTINATARIO BASADA EN UN ESQUEMA DE CONTROL DISTRIBUTED CONTROL SCHEME

En el esquema DCS no existe ningún IDMS Manager, como en el esquema de control SMS. Por tanto, toda la lógica contenida en dicha entidad pasa a integrarse en los propios Clientes IDMS. Por lo tanto, son dichos clientes los que se encargan de calcular la denominada *session async* (que bajo este esquema es la mayor diferencia existente entre los procesos de reproducción de todos los clientes involucrados) y, si esta supera el umbral *session threshold*, deberán calcular también el punto de referencia para la sincronización (*sync reference*). En dicho caso, si la diferencia entre el proceso de reproducción del Cliente IDMS y la *sync reference* (es decir, el valor del *IDMS internal async*), supera el valor máximo de asincronía tolerable, definido como *IDMS internal async threshold*, deberán realizar un ajuste en su proceso de reproducción con el fin de alcanzar un nivel de sincronización IDMS satisfactorio. La Fig. 34 muestra el diagrama de flujo del comportamiento de un Cliente IDMS bajo un esquema de control DCS.

En dicha figura no se muestra cómo se crea una sesión IDMS, ya que ha sido explicado anteriormente. Se puede observar que, tras unirse satisfactoriamente a una sesión IDMS, el Cliente IDMS realiza dos tareas en paralelo. Por un lado, cada $P_{IDMS} \pm \Delta$ segundos, envía un mensaje IDMS Report, donde Δ permite pseudo-aleatorizar el periodo de transmisión con el fin de evitar que todos los clientes IDMS envíen sus informes a la vez y se puedan generar problemas de congestión. Por otro lado, el Cliente IDMS queda en espera de recibir nuevos mensajes IDMS. Nótese que los del tipo *IDMS Hello Request/Response* (explicados anteriormente) no se han incluido en la figura por simplificarla y que sea más fácil de interpretar.

Adicionalmente, la Fig. 35 muestra tanto una situación normal (en la parte izquierda) como una no deseada (en la parte derecha) para los dos esquemas de control propuestos. Se considera una situación normal a aquella en la que todos los Clientes IDMS involucrados envían los mensajes *IDMS Report* dentro del periodo de tiempo definido (*Sync Round*). Cuando el IDMS Manager o un Cliente IDMS (en el esquema SMS o DCS, respectivamente) ha recopilado todos los mensajes *IDMS Report*, calcula el valor de *session async* y comienza una nueva *Sync Round* (cuyo valor es un número que se va incrementando con cada ronda que pasa). Por el contrario, una situación no deseada implica que, al menos, un Cliente IDMS no envíe su mensaje *IDMS Report* dentro del límite establecido por el *Sync Round*. Si esto ocurre, el IDMS Manager (en SMS) o el Cliente IDMS (en DCS) calculará el valor de la *session async* con la información de los Clientes IDMS que haya podido recopilar (siempre y cuando el número de mensajes *IDMS Report* recibidos sean considerados como suficientes como, por ejemplo, el caso en el que 75% de los Clientes IDMS hayan enviado un mensaje *IDMS Report*). En este caso, los mensajes *IDMS Report* que se reciban tarde (es decir, sean de una *Sync Round* anterior a la actual), serán descartados automáticamente.

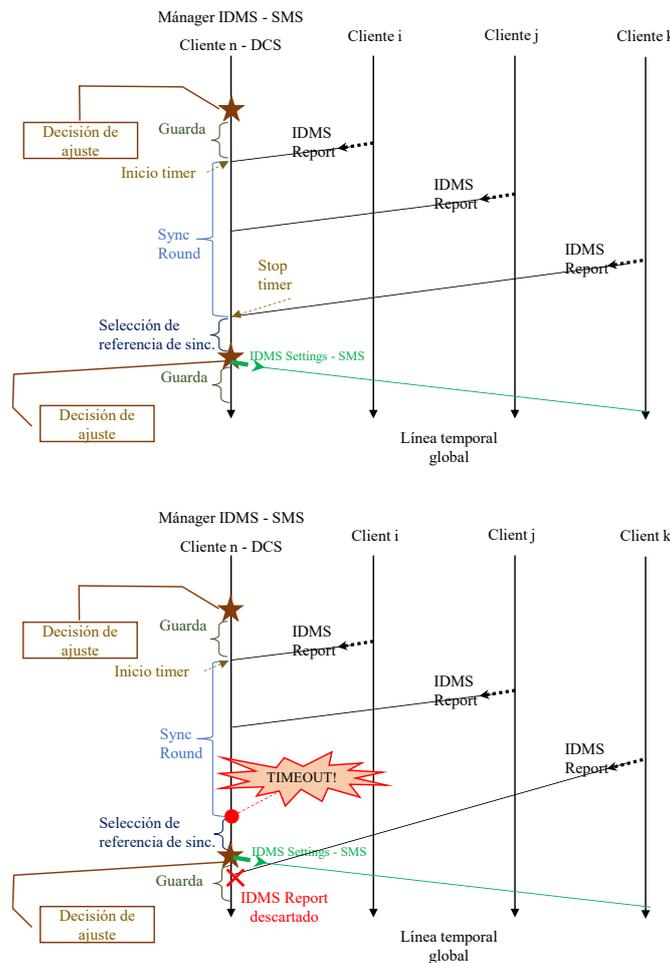


Fig. 35 Situación normal (superior) y no-deseada (inferior)

CÁLCULO DE LA REFERENCIA DE SINCRONIZACIÓN EN UNA SESIÓN DE SINCRONIZACIÓN INTER-DESTINATARIO

Una vez el Mánager IDMS (bajo un esquema de control SMS) o el Cliente IDMS (bajo un esquema de control DCS) ha recibido un número suficiente de mensajes del tipo IDMS Report, lleva a cabo el cálculo de la referencia de sincronización, para ajustar (o no, si no se supera el umbral establecido) el proceso de reproducción a dicha referencia. La referencia de sincronización puede establecerse al Cliente IDMS más lento, al más adelantado, o a un punto medio de los procesos de reproducción de todos los clientes IDMS.

A continuación, se exponen los cálculos para obtener la referencia de sincronización a partir de la información que recibe el Mánager IDMS (bajo un esquema de control SMS) o el Cliente IDMS (bajo un esquema de control DCS), por parte de los Clientes IDMS de una misma sesión.

REFERENCIA DE SINCRONIZACIÓN AL CLIENTE MÁS LENTO

Para calcular la referencia de sincronización al más lento, se itera sobre la información contenida en el mensaje IDMS Report (el instante de generación y el instante de reproducción) de todos los Clientes IDMS y se lleva a cabo el cálculo definido en la Ecuación 2 y en la Ecuación 3, asignando como referencia el instante de generación y de reproducción cuyo valor de asincronía sea el máximo de entre todos los Clientes IDMS.

En el caso de haber adoptado un esquema de control SMS, la información se encuentra contenida en una tabla donde se almacena la información temporal de cada Cliente IDMS (ver Fig. 23). En el caso de haber adoptado un esquema de control DCS, es cada Cliente IDMS el que gestiona una tabla donde se almacena esta misma información.

REFERENCIA DE SINCRONIZACIÓN AL CLIENTE MÁS ADELANTADO

Para calcular la referencia de sincronización al más adelantado, se itera sobre la información contenida en el mensaje IDMS Report (el instante de generación y el instante de reproducción) de todos los Clientes IDMS y se lleva a cabo el cálculo definido en la Ecuación 2 y en la Ecuación 3, asignando como referencia el instante de generación y de reproducción cuyo valor de asincronía sea el menor de entre todos los Clientes IDMS. Se trata de la misma lógica mostrada en la subsección anterior, pero invirtiendo la lógica de decisión para asignar como referencia de sincronización al Cliente IDMS más adelantado.

REFERENCIA DE SINCRONIZACIÓN A UN PUNTO MEDIO DEL PROCESO DE REPRODUCCIÓN

Para calcular la referencia a un punto del proceso de reproducción medio, se adapta la Ecuación 2 para calcular la posición objetivo basada en un valor medio de la diferencia entre el instante de reproducción y de generación de cada uno de los Clientes IDMS de la sesión, tal y como se muestra en la Ecuación 5 (se sigue la misma nomenclatura que la mostrada en la Fig. 20):

Ecuación 5 Cálculo de la referencia de sincronización IDMS a un punto medio del proceso de reproducción

$$\text{Instante de generación objetivo} = \text{instante } v + \left(\frac{\sum_{i=0}^t (\text{instante } m_i - \text{instante } n_i)}{t} \right) \quad (5)$$

Una vez obtenido el instante de generación objetivo, a partir de la Ecuación 5, se emplea la Ecuación 3 para obtener el valor de asincronía para cada uno de los Clientes IDMS de la sesión.

4.7 DISEÑO Y ESPECIFICACIÓN DE UNA SOLUCIÓN DE SINCRONIZACIÓN DE LA REPRODUCCIÓN DE CONTENIDOS HÍBRIDOS Y LA GENERACIÓN DE EFECTOS MULTISENSORIALES

Tal y como se ha expuesto anteriormente, la existencia de efectos multisensoriales (contenidos mulsemmedia) asociados con el contenido principal se señala en el fichero RMCF, al igual que la disponibilidad de contenido audiovisual adicional, la ubicación de la referencia de reloj global, etc. En particular, el contenido o información mulsemmedia se almacena en un fichero SEM (siguiendo el esquema SEDL, definido en MPEG-V [55]), donde se asocia a un instante de un contenido audiovisual un efecto multisensorial indicando su intensidad, duración o cualquier otro tipo de atributo necesario para su ejecución. Sin embargo, el fichero SEM asocia cada efecto multisensorial con un instante de reproducción (PTS timestamp en MPEG-2 TS) del contenido audiovisual. Esto presenta un problema, tal y como se ha expuesto en el Capítulo 3 de trabajos relacionados, puesto que el valor de estas referencias temporales puede ser modificado a lo largo de la cadena de transmisión broadcast. Por tanto, con el fin de que sea posible la sincronización de contenido audiovisual transmitido a través de redes broadcast con contenido mulsemmedia, se propone añadir un atributo más al fichero SEM. Este atributo, denominado *global-timestamp*, contiene el instante temporal basado en NTP en el que se debe generar/activar el efecto durante la presentación del contenido AV (Fig. 36).

```
<sedl:Effect xsi:type="sev:LightType" autoExtraction="both" activate="true" si:anchorElement="true" si:pts="0"
global-timestamp="3763721890965118886"/>
<sedl:GroupOfEffects si:anchorElement="true" si:pts="180000" global-timestamp="3763721892965118886">
  <sedl:Effect xsi:type="sev:ScentType" intensity-value="100.0" intensity-range="0.0 100.0" activate="true"
  sev:scent="forest"/>
  <sedl:Effect xsi:type="sev:TemperatureType" intensity-value="40.0" intensity-range="0.0 40.0"
  activate="true"/>
</sedl:GroupOfEffects>
```

Fig. 36 Fichero SEM con el atributo *global-timestamp*

Cuando se genera la señalización, se parte del contenido AV ya creado. Para poder convertir el valor PTS de las tramas MPEG de dicho contenido en un valor basado en NTP, se extrae del contenido audiovisual el valor NTP insertado en el primer descriptor TEMI *timeline* del contenido (se corresponderá con el valor incluido en el atributo *temi_init* de la etiqueta *MEDIA* correspondiente a ese contenido AV en el RMCF) y se asocia dicho instante al valor PTS = 0 (correspondiente a la primera trama del contenido AV). A continuación, el resto de los valores PTS (en Hz) de cada trama se convierten en formato NTP al convertirlos en unidades de segundos y sumando dicha cantidad de segundos al valor NTP del primer TEMI *timeline* insertado en el contenido AV (Fig. 37).

En el fichero SEM, en el atributo *global-timestamp* se utilizarán los timestamps en NTP obtenidos de esta forma para indicar en qué instante se debe generar/activar cada uno de los efectos multisensoriales.

De esta manera, tanto el contenido audiovisual como el multisensorial comparten la misma línea temporal global una vez han sido transmitidos hasta su destino y es posible alcanzar un nivel de sincronización satisfactorio entre el contenido audiovisual y la generación/activación de los efectos multisensoriales.

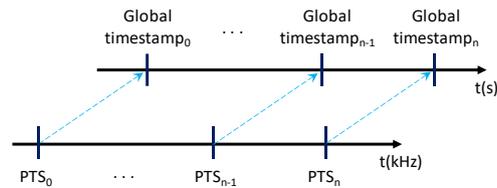


Fig. 37 Proceso de asociación de un valor basado en NTP a partir de un valor PTS

4.7.1 CONTROLADOR MULSEMEDIA

El contenido mulsemmedia puede involucrar diferentes sentidos, por lo que se propone el diseño de un Controlador Mulsemmedia (en adelante MC, de *Mulsemmedia Controller*) para establecer un canal de comunicación entre la app que envíe las acciones asociadas a los efectos sensoriales (ya sea la MS app o una CS app) y los dispositivos encargados de activar/generar dichos efectos. De esta manera, se proporciona una solución modular e independiente de los dispositivos mulsemmedia que pueda haber disponibles, ya que la comunicación siempre se lleva a cabo entre las apps MS o CS y el MC (Fig. 38).

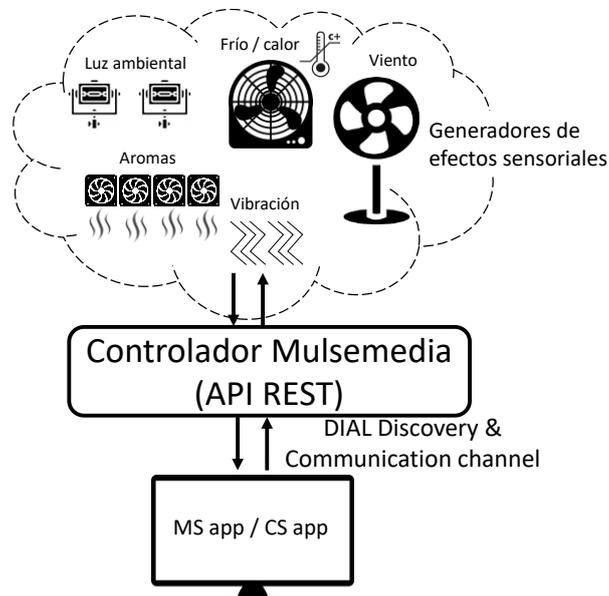


Fig. 38 Descubrimiento del MC a través de DIAL [30]

Respecto a las especificaciones del MC, su principal función consiste en ejercer de pasarela entre la MS o la CS app y los dispositivos que presentan los efectos sensoriales (podría ser implementado en dispositivos microcontroladores de bajo coste, como, por ejemplo, un Arduino). Respecto a la conectividad de dichos dispositivos, se propone emplear tecnologías inalámbricas, puesto que un entorno demasiado cableado puede resultar molesto para el usuario. En caso de ser un dispositivo inalámbrico, se debería implementar un mecanismo de descubrimiento del mismo (Fig. 38), como el descrito en DIAL. En ese caso, al emplear tecnologías inalámbricas, se añade una latencia extra, la cual no debería superar (en condiciones normales, -es decir, sin congestión de la red, sin errores de transmisión, etc.- los 60ms [150]). Teniendo en cuenta que para contenidos sensoriales, una diferencia por debajo de 1 segundo entre la percepción de los efectos y el contenido principal siendo visualizado se considera un nivel de sincronización aceptable [151], dicha latencia extra se considera poco significativa.

Para facilitar el uso de dispositivos con rendimiento reducido (por ejemplo, un microcontrolador), se ha simplificado el mecanismo de comunicación entre dispositivos al máximo posible. En este caso, el comportamiento del MC consiste en recibir órdenes de la MS app o CS para disparar, presentar o terminar determinados efectos sensoriales. Por tanto, tras realizar el

proceso de descubrimiento entre las MS o CS apps y el MC, la MS o CS app envía los mensajes correspondientes a efectos multimedia a una API (*Application Programming Interface*) REST (*Representational State Transfer*), explicada a continuación, establecida en el MC. De esta forma, a partir de peticiones HTTP, como GET, POST o DELETE, las MS o CS apps pueden enviar órdenes al MC de una manera directa y simple.

4.7.2 USO DE UNA INTERFAZ DE PROGRAMACIÓN DE APLICACIONES REST

En esta sección se va a describir el funcionamiento de una API REST y cómo ha sido adaptada en la Tesis Doctoral para permitir el envío de mensajes de la MS o CS app al MC.

En particular, una API REST tiene las siguientes características principales:

- Arquitectura cliente/servidor sin estado: cada petición HTTP contiene toda la información necesaria para ejecutar una acción. Al tratarse de peticiones HTTP, cuyas operaciones disponibles son GET, POST, PUT y DELETE, se proporciona una interfaz uniforme.
- Utiliza la URI como identificador único de cada recurso disponible.
- Es independiente del tipo de lenguaje o plataformas utilizadas.

En el caso específico de la presente Tesis Doctoral, se definen dos tipos de acciones diferentes. Por una parte, la petición *HTTP GET* se emplea para que las MS o CS apps puedan obtener, del MC, los diferentes tipos de efectos sensoriales disponibles. Por otro lado, la petición *HTTP POST* se utilizará para indicar al MC que se debe disparar o terminar un efecto sensorial. A continuación, la Tabla 6 muestra cada una de las acciones definidas.

Tabla 6 Tipos de acción del MS/CS app al MC

Tipo de acción	Estructura URL	Descripción de la respuesta
GET	http://IP_MC/list_available_effects	El MC devuelve una lista de los efectos sensoriales disponibles.
POST	http://IP_MC/set/{efecto}/[on, off, {intensidad}]	Indica al MC el estado en el que debe poner un determinado efecto, ya sea un valor binario (por ejemplo, encender o apagar un relé que activa un ventilador) o un valor analógico (por ejemplo, encender un ventilador a una determinada intensidad).

En la Fig. 39 se muestra un ejemplo en el que se realiza una petición a la API REST para interactuar con un recurso disponible (superior) o no disponible (inferior).

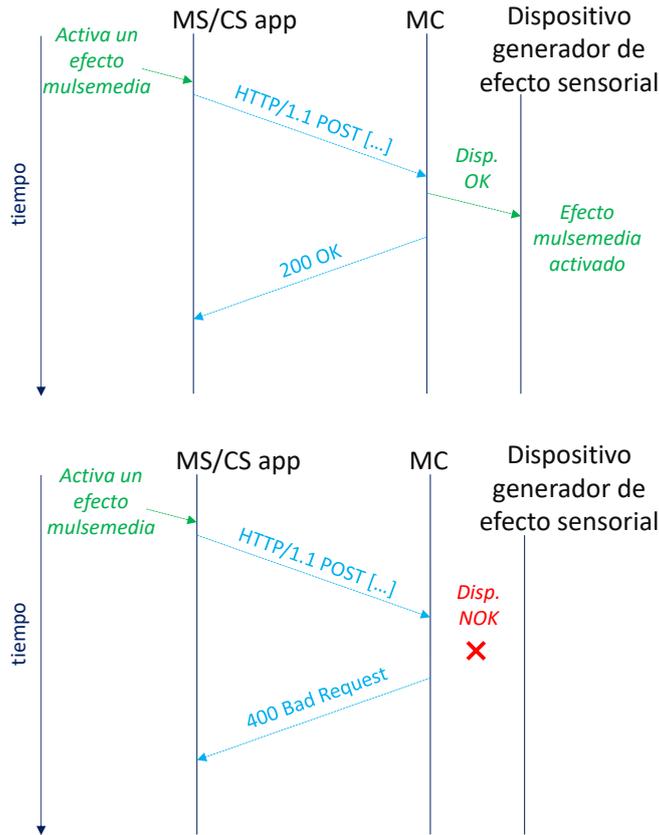


Fig. 39 Ejemplo de una petición a un recurso disponible (izquierda) o no disponible (derecha) entre la MS/CS app y el MC

4.7.3 GESTIÓN Y SINCRONIZACIÓN DEL CONTENIDO MULSEMEDIA

Tal y como se ha expuesto en secciones anteriores, una vez el MS recibe el contenido broadcast, a partir del descriptor TEMI *location* puede obtener el fichero RMCF conteniendo la información adicional respecto a los contenidos complementarios al flujo audiovisual principal, incluida la URL del fichero SEM, que permite asociar los contenidos multisensoriales con el contenido audiovisual. La MS app o la CS app analizarán este fichero para gestionar los diferentes efectos sensoriales que pueden ser generados durante la reproducción del contenido audiovisual. Dichas apps, a partir del fichero SEM, crean una tabla (Tabla 7) en la que se incluye:

- La marca temporal global (atributo *global-timestamp*) asociada para cada efecto sensorial.
- El tipo de efecto sensorial que debe ser gestionado.
- Otra información necesaria, como la duración e intensidad del efecto, el tipo (por ejemplo, el tipo de olor si es un efecto de aroma), o cualquier otro tipo de información, separadas por punto y coma, dependiendo del tipo de efecto.

Tabla 7 Ejemplo de tabla de efectos mulsemmedia almacenada en la MS/CS app

Timestamp asociado (ns)	Tipo de efecto	Información adicional
3763721890965118886	Aromas	tipo=pino; intensidad=100/100
3763721892965118886	Viento	intensidad=50/100
3763721896965118886	Vibración	intensidad=10/100
[...]	[...]	[...]

Por tanto, como la MS app/CS es consciente del instante de tiempo global del contenido que está presentando y, además, conoce el instante de tiempo global en el que deben ser activados los diferentes efectos sensoriales, solamente debe comparar ambos valores para decidir cuándo transmitir la orden de ejecución de un determinado efecto sensorial al MC. En la Ecuación 6, se muestra el cálculo del instante en el que se debe enviar la orden de activación de un determinado efecto multisensorial.

Ecuación 6 Cálculo del instante de activación de los efectos multisensoriales

$$t_{activación} = t_{contenido} - \frac{RTT_{MS/CS-MC}}{2} - \Delta t_{MC} \quad (5)$$

Siendo $t_{contenido}$ el instante global del contenido asociado para el que debe presentarse el efecto multisensorial; $RTT_{MS/CS-MC}$ la latencia asociada (*Round-Trip Time*) al envío y respuesta de mensajes entre el MS o el CS app al MC; y Δt_{MC} el tiempo asociado a la latencia desde que el MC recibe la orden de activación hasta que dicha activación se hace efectiva. La Fig. 40 muestra el diagrama de flujo de las diferentes tareas involucradas en la gestión de los efectos sensoriales por parte de la MS o CS app y de su transmisión al MC.

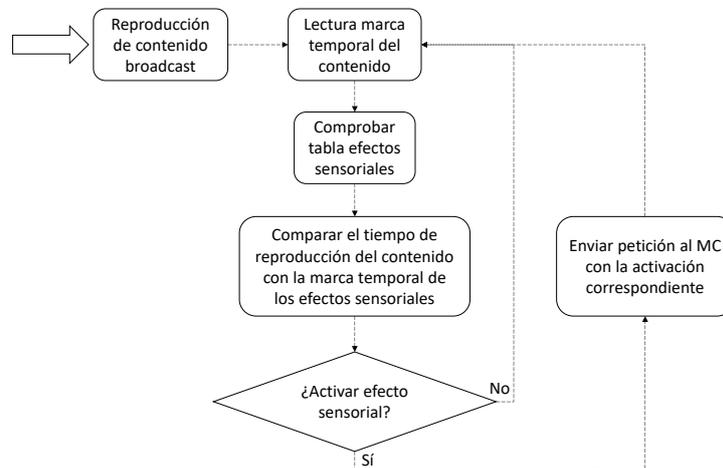


Fig. 40 Diagrama de flujo de las principales tareas para la activación de contenidos multimedia por parte de las apps MS o CS

4.8 CONCLUSIONES

En este capítulo se ha descrito el conjunto de mecanismos de sincronización híbrida propuestos para posibilitar nuevas experiencias televisivas enriquecidas, personalizadas, inmersivas y compartidas. Previamente, se han descrito los requisitos necesarios para que se puedan establecer satisfactoriamente estos mecanismos de sincronización, como son los siguientes: 1) los contenidos híbridos deben estar relacionados de alguna manera, con el fin, por ejemplo, de poder localizar contenido complementario al contenido principal; 2) los procesos, dispositivos o destinatarios (según escenario) deben descubrirse y deben establecer algún tipo de canal bidireccional que les permita comunicarse entre sí; 3) los procesos, dispositivos o destinatarios involucrados en la misma sesión de consumo deben soportar y reproducir contenido recibido a través de la red broadcast y también a través de la red broadband; y 4) los procesos de reproducción involucrados deben soportar mecanismos de ajuste del estado de reproducción con el fin de alcanzar un nivel de sincronización satisfactorio.

Capítulo 5. PLATAFORMA EXTREMO-A-EXTREMO PARA LA GENERACIÓN, DISTRIBUCIÓN Y CONSUMO SINCRONIZADO DE CONTENIDOS HÍBRIDOS, INCLUYENDO CONTENIDOS OMNIDIRECCIONALES Y EFECTOS MULTISENSORIALES

En este capítulo se describe la plataforma extremo-a-extremo desarrollada en el seno de la presente Tesis Doctoral, que incluye diferentes módulos para la generación, distribución y consumo sincronizado de contenidos híbridos, incluyendo contenidos omnidireccionales y efectos multisensoriales. Dicha plataforma constituye una de las contribuciones más relevantes de la Tesis Doctoral y se ha utilizado como testbed para implementar en ella y validar las soluciones de sincronización definidas en la misma, descritas en el capítulo anterior.

5.1 INTRODUCCIÓN

En las siguientes secciones de este capítulo se describe la plataforma desarrollada en esta Tesis Doctoral, de manera desglosada y gradual. Se expone su arquitectura, los módulos que la componen así como todas las funcionalidades contenidas en los mismos (como, por ejemplo, los mecanismos de sincronización IDES e IDMS propuestos, así como la generación y reproducción de contenidos tradicionales, omnidireccionales y multisensoriales. Cabe resaltar que, para cualquiera de los escenarios descritos en el capítulo anterior (IDES, IDMS o Mulsemedia), la plataforma se puede dividir en dos partes claramente diferenciadas: la parte correspondiente al proveedor de contenidos (es decir, la generación, preparación y señalización de los contenidos) y la parte correspondiente al consumidor final del contenido (es decir, la recepción, análisis, adquisición y reproducción de contenidos, así como los mecanismos de sincronización necesarios).

5.2 ARQUITECTURA DE LA PLATAFORMA

Tal y como se observa en la Fig. 41, la plataforma desarrollada está comprendida de las dos partes mencionadas bien diferenciadas: la parte del proveedor de contenidos y la parte del usuario o consumidor final.

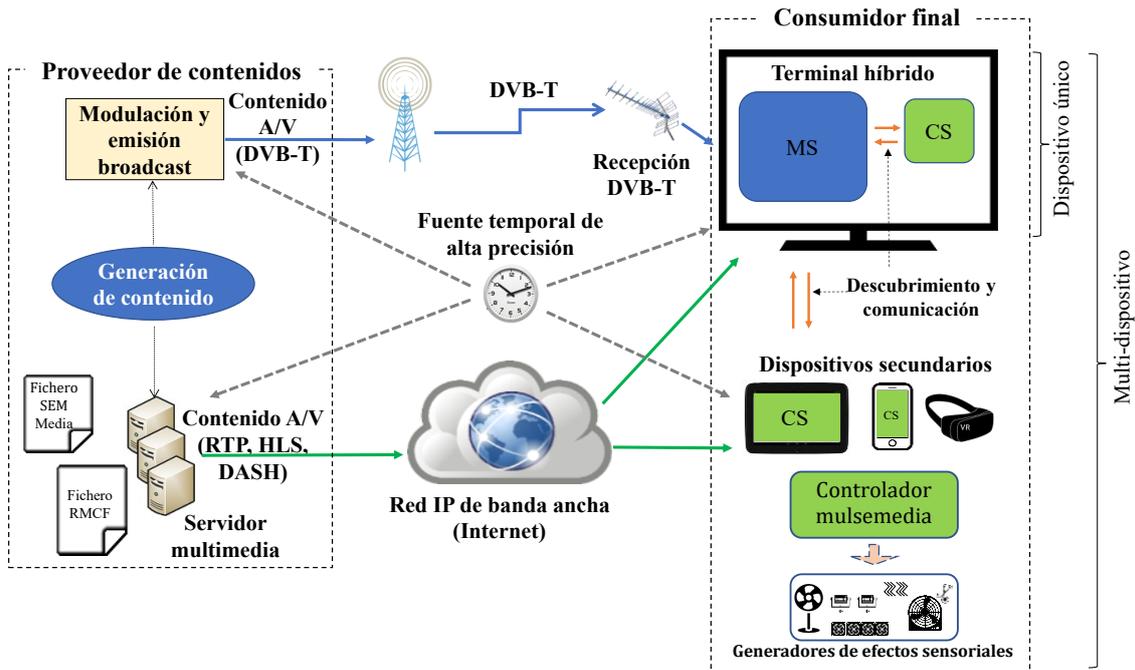


Fig. 41 Visión general/arquitectura de la plataforma desarrollada.

En la parte correspondiente del proveedor de contenidos, se ubican los módulos asociados a la generación audiovisual del contenido, así como de su codificación en diferentes formatos para poder ser transmitido a través de distintas tecnologías de transmisión (por ejemplo, DVB-T e Internet). Adicionalmente, en esta parte de la plataforma se incluyen los elementos necesarios para que el consumidor final pueda acceder a toda la información necesaria con el fin de poder adquirir y reproducir contenido complementario o bien adquirir la información asociada a los efectos multisensoriales disponibles en el contenido. Por tanto, se puede deducir que, en la parte del proveedor de contenidos, será necesaria la existencia de uno o varios servidores multimedia, encargado de almacenar tanto el contenido disponible a través de Internet, como de los ficheros que almacenan la información asociada. La Fig. 42 muestra un resumen de las funcionalidades incluidas en esta parte de la plataforma.

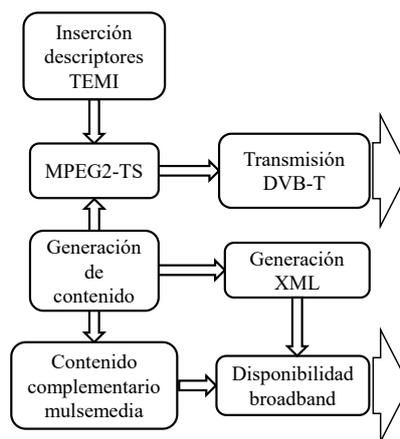


Fig. 42 Funcionalidades incluidas en la parte del proveedor de contenidos.

En la parte correspondiente al consumidor final, se incluyen los módulos encargados de recibir el contenido e interpretar la información asociada al mismo, con el objetivo de reproducir de forma sincronizada, el contenido principal y el contenido complementario, ya sea contenido multimedia tradicional, omnidireccional o multimedial. Para ello, es necesario establecer

mecanismos de descubrimiento y comunicación entre los dispositivos involucrados (excepto si solamente hay un único dispositivo -terminal híbrido-, el cual puede reproducir el contenido principal y el contenido complementario sin necesidad de llevar a cabo estas tareas).

Respecto al mecanismo de sincronización IDEs, se trata de un módulo que está ubicado en la parte del consumidor final, y hace uso de una comunicación entre dispositivos que se encuentran en una misma red IP local. Sin embargo, cuando se trata del mecanismo de sincronización IDMS. Es necesario una gestión de la sincronización entre destinatarios que puede estar ubicada en diferentes partes, según el esquema de control de sincronización adoptado. Tal y como se puede observar en la Fig. 43, si se adopta un esquema de control SMS, el módulo encargado de gestionar la sincronización entre destinatarios (IDMS Manager) debe ser accesible por todos los consumidores involucrados (por sus clientes IDMS) a través de un canal de comunicación bidireccional (esto es, Internet). Si se adopta un esquema de control distribuido DCS, los módulos de sincronización IDMS (Clientes IDMS) se encuentran en los respectivos terminales involucrados.

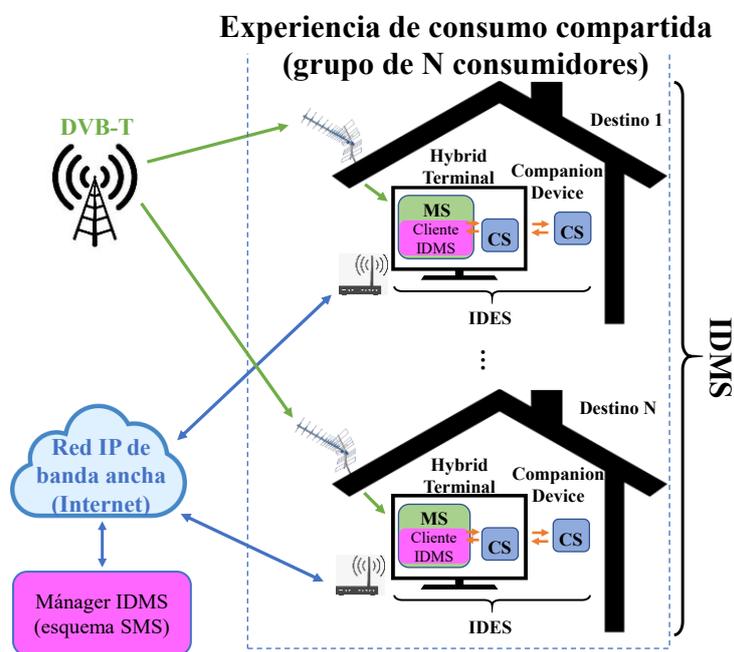


Fig. 43 Parte del consumidor final en un escenario IDMS.

Como resumen, las Fig. 44 y Fig. 45 muestran los módulos incluidos en la MS app, en la parte del consumidor final de la plataforma, incluyendo los módulos necesarios en el escenario IDMS e IDEs, respectivamente. La Fig. 46 muestra los módulos incluidos en la CS app, en la parte del consumidor final de la plataforma, incluyendo módulos para los mecanismos IDEs necesarios.

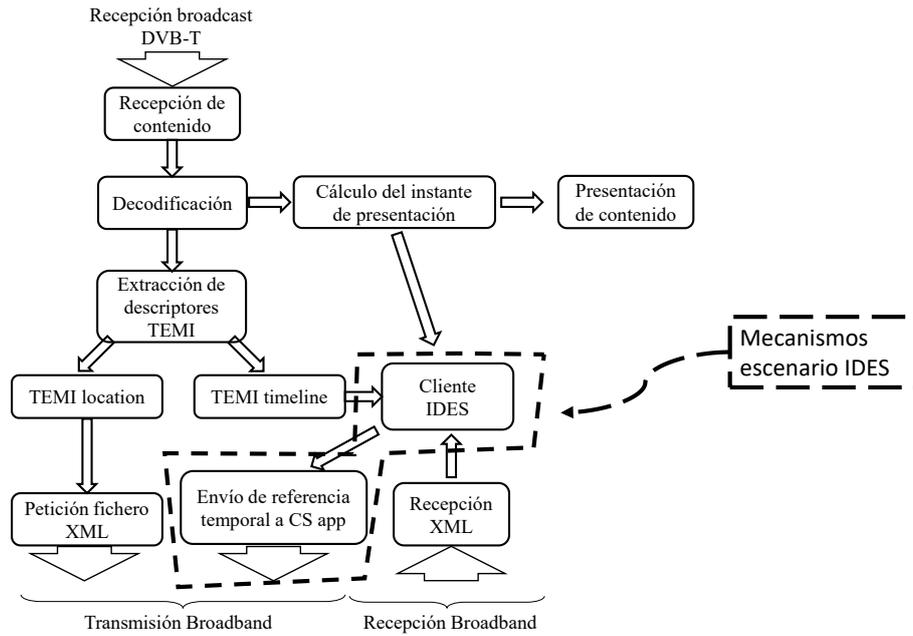


Fig. 44 Funcionalidades involucradas en la parte del consumidor final (MS app), incluyendo los módulos necesarios para un escenario IDES.

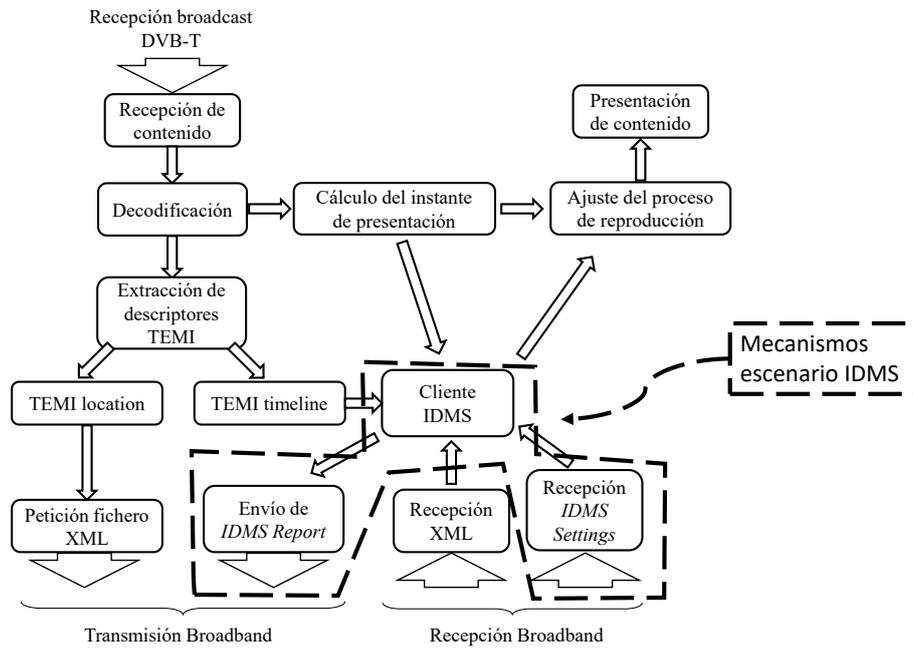


Fig. 45 Funcionalidades involucradas en la parte del consumidor final (MS app), incluyendo los módulos necesarios para un escenario IDMS.

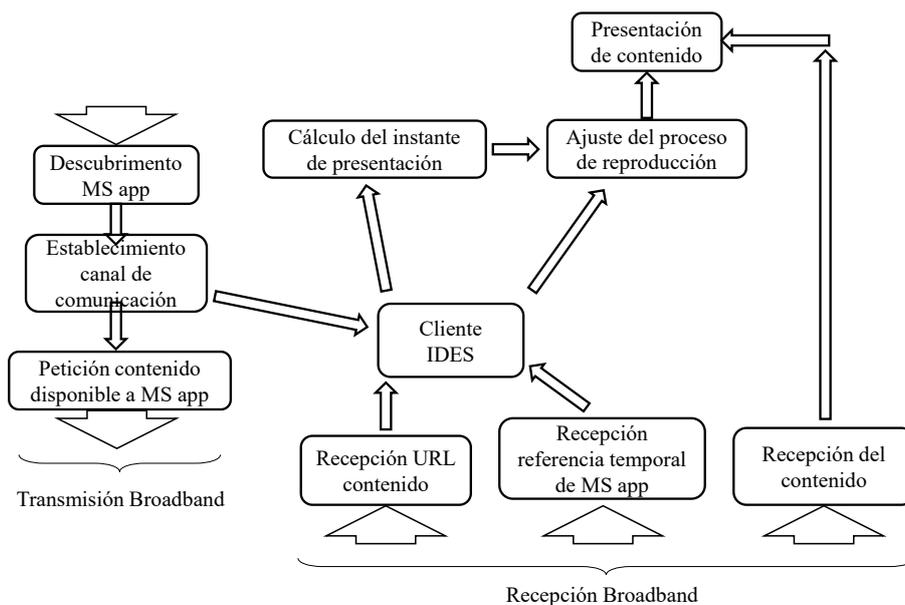


Fig. 46 Funcionalidades involucradas en la parte del consumidor final (CS app), incluyendo los módulos necesarios para un escenario IDES.

En los apartados 5.4 y 5.5 se explican con más detalle las dos partes de la plataforma (proveedor de contenidos y consumidor final, respectivamente). En los siguientes subapartados se describen otros módulos adicionales que, por su especial relevancia, se exponen en sus propias subsecciones como, por ejemplo, los módulos de sincronización multimedia (IDES e IDMS) o porque son módulos comunes a ambas partes (proveedor de contenidos y consumidor final) como, por ejemplo, el módulo de sincronización de relojes y referencia temporal global.

5.2.1 MÓDULO DE SINCRONIZACIÓN DE RELOJES Y REFERENCIA TEMPORAL GLOBAL

La plataforma también integra un módulo servidor de reloj global para la inserción de marcas temporales extrínsecas en todos los contenidos disponibles (tanto broadcast como broadband). Esto permite que los mecanismos de sincronización implementados en la parte del consumidor final puedan establecer un nivel de sincronización satisfactorio entre los procesos de reproducción de contenidos híbridos que están siendo consumidos de manera simultánea en uno o varios dispositivos (en el caso de un escenario IDES) o localizaciones (en el caso de un escenario IDMS). Este módulo consiste en un servidor NTP. Se ha utilizado el servidor NTP perteneciente a la Entidad Pública Empresarial Red.es, ubicado en Valencia (IPv4 147.156.7.50).

5.2.2 MÓDULOS DE SINCRONIZACIÓN MULTIMEDIA

En cuanto a los contenidos AV, por un lado, en la plataforma la sincronización intra- e inter-flujo (en cuanto a los flujos incluidos en el múltiplex MPEG2-TS) está garantizada por el framework GStreamer, utilizado para su implementación. GStreamer es un marco de desarrollo libre, multi-plataforma, que permite crear aplicaciones multimedia, tanto de reproducción como de manipulación de contenido. A continuación, se explican brevemente sus características y elementos principales, así como los conceptos necesarios para comprender su funcionamiento.

Está basado en la configuración de *pipelines* (tuberías) para desarrollar aplicaciones multimedia. Una *pipeline* es un contenedor de nivel superior que proporciona un bus de comunicaciones secuencial para todos los componentes o elementos incluidos en la aplicación y que gestiona la sincronización del flujo de ejecución. Cada *pipeline* puede contener una serie de elementos o componentes (definidos en los *plugins*) que, encadenados y de manera combinada, pueden realizar funcionalidades más avanzadas, como las que se llevan a cabo en esta Tesis Doctoral.

A partir de este marco de desarrollo, en la plataforma se han implementado las soluciones de sincronización de la reproducción de contenidos híbridos entre dispositivos (IDES) y entre destinatarios (IDMS) descritas en el Capítulo 4.

SINCRONIZACIÓN INTER-DISPOSITIVO

El esquema de control de sincronización adoptado para IDES en la plataforma es el M/S, donde la MS app adopta el rol de máster y la CS app el rol de esclavo.

Una vez el canal de comunicación App2App entre las MS y CS apps ya está establecido y que la CS app conoce la ubicación del contenido broadband disponible, los pasos necesarios para alcanzar un nivel de sincronización satisfactorio son los siguientes: 1) la MS app extrae la información almacenada en el descriptor TEMI *timeline*, que indica el instante de generación para cada fotograma; 2) la MS app monitoriza el paso de cada fotograma por todos los elementos del proceso de reproducción (recepción, demultiplexación y decodificación) hasta su renderizado; y 3) la MS app estima o calcula el instante de presentación (en la línea temporal del reloj global) de cada trama. Una vez la MS app conoce el instante de generación y el instante de presentación asociado a una trama, envía dicho par de instantes temporales a la CS app a través del canal App2App. El hecho de estimar el instante de presentación, en lugar de utilizar el instante de recepción de contenido permite alcanzar valores más precisos de sincronización, puesto que de esta manera se toman en consideración posibles eventos que añadan una latencia o retardo extra en los diferentes módulos involucrados. Una vez la CS app recibe este par de marcas temporales (instantes de generación y presentación), puede calcular el nivel de sincronización que tiene respecto a la MS app y valorar la realización de un ajuste en el proceso de reproducción. Dicho ajuste se llevará a cabo si se supera un umbral de asincronía máxima tolerable con el proceso de reproducción de la MS app, cuyo valor también es configurable.

En la plataforma desarrollada, se han implementado dos tipos de ajustes en el proceso de reproducción: bruscos (S&P) o suaves (AMP), según la magnitud de la asincronía obtenida para un instante determinado. Si la asincronía supera el umbral de asincronía máxima tolerable, pero se encuentra por debajo de un valor equivalente a dos veces dicho umbral (en adelante, este valor se denominará *umbral superior de asincronía*), se aplicarán técnicas de ajuste AMP, consistentes en modificar la tasa de reproducción a un valor por encima o por debajo de la tasa nominal (por ejemplo, al 80% o al 120% de su velocidad normal). En el caso de que el valor de asincronía obtenido también supere el umbral superior de asincronía, se realizará un ajuste brusco en el proceso de reproducción mediante técnicas de S&P, ya que permiten corregir de una forma más rápida (aunque también abrupta y, por tanto, más perceptible para el usuario) el nivel de sincronización.

SINCRONIZACIÓN INTER-DESTINATARIO

En el caso específico de un escenario IDMS con un esquema de control SMS, es necesaria la figura de la entidad denominada IDMS Manager. En la implementación de esta plataforma, dicha entidad consiste en un servidor WebSocket (WS) encargado de gestionar las diferentes sesiones IDMS existentes, además de calcular periódicamente la referencia de sincronización para cada una de las sesiones IDMS activas en la plataforma y enviarla a los clientes IDMS de cada sesión.

Por el contrario, en el caso de un escenario IDMS con un esquema de control distribuido DCS, son los propios Clientes IDMS los responsables de realizar las acciones para poder alcanzar un estado de sincronización satisfactorio a través del intercambio y procesado de los mensajes definidos en el capítulo anterior. Este intercambio de mensajes entre Clientes IDMS se lleva a cabo a través de una dirección IP multicast, tal y como se ha expuesto en Capítulo 4.6.

Adicionalmente, se ha integrado el cálculo para obtener la referencia al Cliente IDMS más lento y a un punto del proceso de reproducción medio (ver Capítulo 4.6).

5.2.3 SERVICIO DE CHAT SINCRONIZADO EN EL ESCENARIO DE SINCRONIZACIÓN INTER-DESTINATARIO

Adicionalmente, en la plataforma también se ha implementado un canal de intercambio de emensajes de texto (chat) entre destinatarios con el fin de permitir una experiencia de consumo compartida. Para ello, en la plataforma se ha habilitado un servidor WS adicional para gestionar el servicio de chat. Tal y como se puede observar en la Fig. 47, tanto el IDMS Manager como el servidor de chat, pueden compartir información acerca de las sesiones existentes, así como de los integrantes de cada sesión.

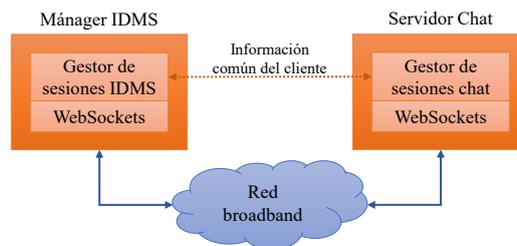


Fig. 47 Relación entre el IDMS Manager y el servidor de chat en un escenario IDMS

Este servicio de chat puede ser utilizado tanto en el dispositivo principal (terminal híbrido) como en los dispositivos secundarios. En el caso de que la herramienta se utilice en el dispositivo secundario, previamente se debe haber ejecutado el proceso de descubrimiento e intercambio de información entre ambos dispositivos, tal y como se ha explicado con anterioridad, a través de los protocolos definidos en DIAL [30]. Además, en el RMCF se ha añadido la información necesaria respecto a la existencia (y ubicación) del servidor de chat existente. La propiedad añadida al RMCF, así como sus atributos, se exponen en la Tabla 8.

Tabla 8 Etiqueta y propiedades integradas en el fichero RMCF para anunciar la disponibilidad de un servicio de chat

Etiqueta	Propiedad	Descripción
CHAT		Especifica la información necesaria para acceder al servidor de chat que haya disponible
	<i>id</i>	Identificador único.
	<i>protocol</i>	Protocolo utilizado (por ejemplo, WS).
	<i>metadata</i>	Breve descripción del servicio disponible.
	<i>uri</i>	Ubicación del servidor de chat.

Con el objetivo de que la herramienta de chat no presente comentarios o mensajes de los participantes, a aquellos usuarios que todavía no hayan visualizado el contenido sobre el que va dirigido dicho mensaje, se ha asignado como metadato para cada mensaje, una marca temporal relacionada con la marca temporal insertada en el descriptor TEMI *timeline* del flujo de contenido MPEG2-TS. De esta forma, se puede asociar el mensaje de chat enviado (o recibido) a un instante específico del contenido y, como consecuencia, la herramienta de chat es capaz de decidir si debe mostrar el mensaje de manera inmediata o bien debe esperar un determinado tiempo (chat sincronizado). Respecto a este mecanismo "anti-spoiler", cabe destacar que no se aplica para los usuarios cuyo instante de reproducción del contenido vaya por delante. En estos casos, si un usuario que esté reproduciendo un instante de contenido de manera atrasada respecto al resto, y envía un mensaje, dicho mensaje se mostrará inmediatamente para todos los usuarios involucrados. Sin embargo, este mecanismo (y estas situaciones) se pueden evitar si se habilita el mecanismo de sincronización IDMS que se ha presentado en esta Tesis Doctoral. El mecanismo integrado en la herramienta de chat presenta una capa más de robustez en cuanto a la

sincronización para no afectar a la QoE percibida por los usuarios, en el caso de que el mecanismo de sincronización IDMS sufra algún fallo eventual.

INTERFAZ DE MENSAJES POR EVENTO Y CONTEXTUALIZADOS

Con el objetivo de poder ofrecer al usuario una forma más rápida e interactiva de comunicarse a través de la herramienta de chat, se ha definido una interfaz de mensajes contextualizados según el género o tipo de contenido consumido en la sesión IDMS. Esto permite que el consumidor experimente y proporcione una conversación más fluida, puesto que este tipo de mensajes predefinidos se refieren y reflejan acciones y eventos que pueden ocurrir en el contenido que se está visualizando. Por ejemplo, para una sesión compartida para ver un partido de fútbol, los mensajes predefinidos que pueden integrarse en esta interfaz podrían ser frases como “¡Golazo!”, “¡Eso es falta!” o “¡Vamos, equipo!”, puesto que, en un partido de fútbol, este tipo de comentarios son previsibles y esperables. Estos mensajes se almacenan en un fichero (por ejemplo, en formato JSON o XML) al que la herramienta de chat tiene acceso para poder personalizar esta interfaz según el contenido visualizado de manera dinámica y transparente para el usuario. En esta Tesis Doctoral, se han almacenado en un fichero XML, mediante las etiquetas y propiedades que se describen en la Tabla 9.

Tabla 9 Etiqueta y propiedades definidas para proporcionar un interfaz de mensajes contextualizados en la herramienta de chat

Etiqueta	Propiedad	Descripción
	<i>id</i>	Identificador único.
BUTTON	<i>btn_name</i>	Nombre del botón.
	<i>msg2send</i>	Mensaje que debe ser enviado una vez se pulse el botón.

En las siguientes secciones, se describen los procesos y funcionalidades de cada uno de los módulos involucrados en la arquitectura de la plataforma desarrollada.

5.3 PARTE DEL PROVEEDOR DE CONTENIDOS

En la parte correspondiente al proveedor de contenidos (Fig. 41, izquierda), la plataforma incluye una serie de módulos encargados de la codificación, encapsulación, segmentación, almacenamiento, modulación y transmisión del contenido multimedia. Además, incluye los módulos encargados de la generación e inserción (o almacenamiento) de metadatos relevantes (por ejemplo, las marcas temporales globales, la descripción y localización del contenido multimedia disponible, etc.).

5.3.1 MÓDULOS DE GENERACIÓN DE CONTENIDOS

La plataforma se ha desarrollado para poder transmitir contenido a través de tecnología DVB-T (broadcast) y tecnologías como HLS, RTP o DASH, basadas en la transmisión a través de redes broadband. La Fig. 42 resume las funcionalidades involucradas en esta parte de la plataforma.

MÓDULO DE GENERACIÓN DE CONTENIDO BROADCAST E INSERCIÓN DE DESCRIPTORES TIMING AND EXTERNAL MEDIA INFORMATION

Por un lado, en el módulo de generación del contenido broadcast, en primer lugar, si es necesario, se hace uso de la herramienta *ffmpeg* [152] para configurar el contenido de vídeo con una tasa de 25 fotogramas por segundo (fps), que es cómo se retransmite actualmente el contenido broadcast en España (TDT), herencia del sistema de codificación analógico PAL [153].

Por otro lado, en dicho módulo también se hace uso del framework GPAC [63] para la encapsulación del contenido multimedia en flujos MPEG2-TS, incluyendo la señalización sobre la existencia de contenidos broadband. Dicha señalización incluye la ubicación del fichero RMCF y las marcas temporales de un reloj global en los descriptores TEMI *location* y *timeline*, respectivamente, explicados en la sección 3 del Capítulo 4. Para ello se utiliza la herramienta MP42TS, incluida en dicho framework.

MÓDULO DE GENERACIÓN DE CONTENIDO BROADBAND

En cuanto a contenidos broadband, en la presente Tesis Doctoral sólo se describe la utilización de la tecnología de transmisión adaptativa MPEG-DASH, ya que es la adoptada por el estándar HbbTV. No obstante, la plataforma también permite la distribución de contenido vía tecnologías broadband mediante HLS [154] y RTSP [155] + RTP/RTCP [156].

Cabe resaltar que las tecnologías adaptativas establecen una conexión unicast con cada uno de los clientes que solicitan la visualización del contenido audiovisual. Es por esto por lo que, además, el contenido está disponible a través de RTSP + RTP/RTCP, ya que de esta manera sí se puede transmitir el contenido a una dirección multicast, ahorrando así en términos de ancho de banda. En particular, RTP se adopta para la transmisión del contenido, RTCP permite intercambiar mensajes de control entre cliente/servidor y RTSP se utiliza para el control y gestión de la sesión del consumo multimedia.

Todo el contenido generado y disponible a través de redes broadband está almacenado en servidores web convencionales a excepción del contenido gestionado a través de RTSP + RTP/RTCP, para el que sí es necesaria la existencia de un servidor específico RTSP.

CONTENIDO AUDIOVISUAL

Para cada uno de los contenidos complementarios de vídeo, tanto tradicional (es decir, 2D) como omnidireccional, se generarán, para cada contenido multimedia, diferentes calidades para su consumo a través de tecnologías adaptativas, como HLS y DASH. En particular, para el contenido basado en DASH, se hace uso de la herramienta descrita en [157], a través de la cual se puede configurar una serie de parámetros tales como la duración de los segmentos o el número de diferentes calidades disponibles para el contenido.

PARTICULARIDADES PARA CONTENIDO OMNIDIRECCIONAL

En la plataforma desarrollada, se ha proporcionado soporte para dos tipos de proyecciones, concretamente la ERP y la CMP, definidas en el Capítulo 3. Por tanto, con el fin de notificar al usuario acerca de la disponibilidad de contenido omnidireccional y, además, de las características y parámetros de dicho contenido, se han definido una serie de propiedades y parámetros adicionales que se integran en el fichero RMCF y se presentan en la Tabla 10.

Tabla 10 Etiquetas del RMCF para soportar contenido omnidireccional

Etiqueta	Propiedad	Descripción
		Especifica los metadatos necesarios para cualquier contenido AV disponible
MEDIA	id	Identificador único
	media_type	Tipo de contenido multimedia (para contenido omnidireccional su valor es '360AV')
	media_format	Información sobre la codificación o el formato (por ejemplo, H264)
	metadata	Breve descripción
	temi_init	Tiempo global absoluto cuando se generó el inicio del contenido
		Permite especificar diferentes fuentes para el mismo contenido
SOURCE	projection	Tipo de proyección utilizada (ERP o CMP)
	protocol	Protocolo utilizado (por ejemplo, HTTP)
	tiled	Booleano utilizado para indicar si se trata de contenido dividido en <i>tiles</i> o no
	uri	Identificador único del contenido omnidireccional

Tal y como se puede observar, para identificar la disponibilidad de contenido omnidireccional se debe indicar en el tipo de contenido de la etiqueta MEDIA. Además, se ha considerado añadir otras propiedades tales como *projection* o *tiled*, que indiquen el tipo de proyección utilizada o si se trata de contenido codificado en *tiles*, respectivamente, aunque en la versión actual de la plataforma dichas propiedades no se encuentran implementadas. En el caso de tratarse de contenido dividido en *tiles* y no contar con dispositivos compatibles con esta tecnología, el contenido pasará a solicitarse a través del nodo de procesado en la nube, cuya ubicación está reflejada en la propiedad *uri*. De esta manera, es el nodo quien asumirá el coste del procesado del contenido basado en *tiles* y, de cara al usuario o consumidor final, todo este proceso será transparente y permitirá el consumo de contenido omnidireccional sea cual sea el dispositivo secundario utilizado.

Respecto a la generación de contenido omnidireccional disponible a través de la red broadband, no existe ninguna diferenciación en el procedimiento del uso de las herramientas utilizadas para la generación del contenido 2D convencional, salvo que las resoluciones y bitrates máximos y mínimos necesarios que deben ser generados para este tipo de contenidos son significativamente mayores que para el contenido convencional.

CONTENIDO MULSEMEDIA

Respecto a la generación del contenido mulsemmedia, se ha adoptado la herramienta SeVino2, descrita en la sección 0 ([121]). Con esta herramienta, se puede generar de forma gráfica e intuitiva cualquier tipo de efecto sensorial para complementar el contenido audiovisual, a través de un reproductor integrado sobre una serie de etiquetas que pueden insertarse para posteriormente generar el fichero SEM donde estará toda la información multisensorial contenida. Como ejemplo, en la Fig. 48 se muestra una captura de cómo insertar un efecto de viento con una intensidad 75/100 durante 4 segundos:

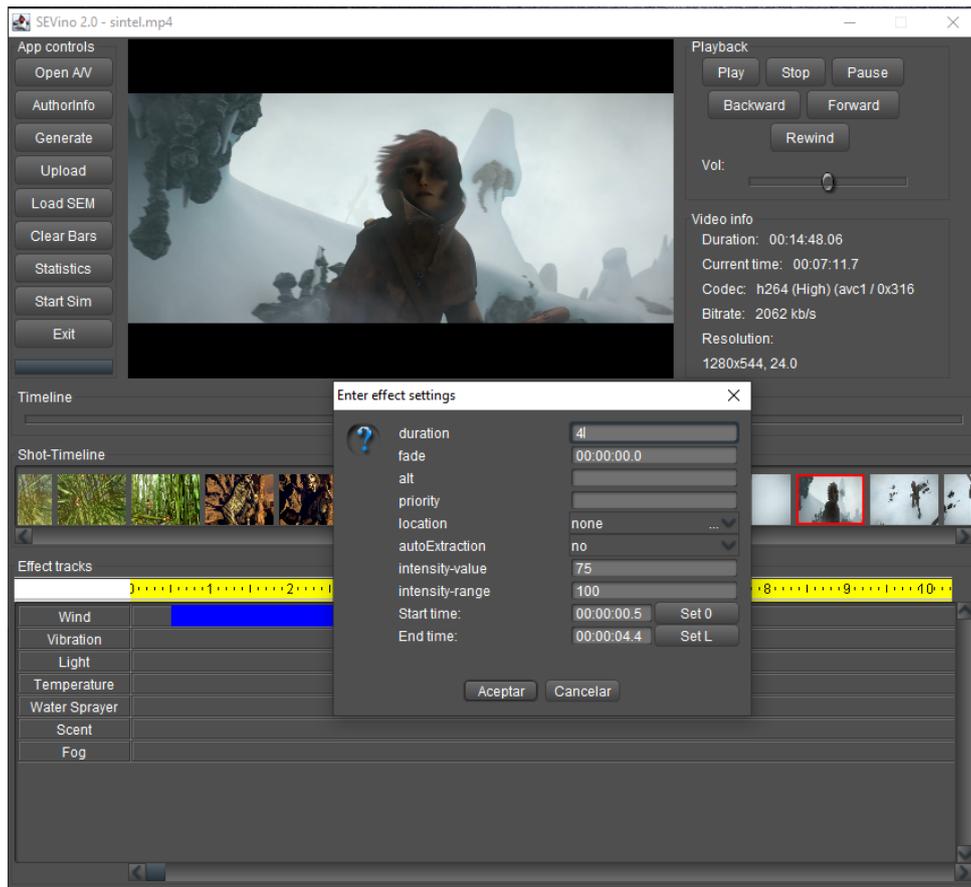


Fig. 48 Herramienta SeVino2 para la inserción de efectos multisensoriales

Una vez se ha etiquetado con la herramienta SeVino todos los efectos multisensoriales a través de la interfaz de usuario, se genera un fichero con formato XML donde queda reflejado el tipo de efecto, su instante inicial y final, la intensidad del efecto y el rango de valores aceptados para dicho efecto. Para la generación de este tipo de contenidos, a través de esta herramienta, no es necesario ningún tipo de comando, ya que todo se lleva a cabo a través de la interfaz de usuario disponible.

En esta parte de la plataforma se genera el contenido mulsemmedia. A diferencia de las anteriores subsecciones (las relacionadas con contenido multimedia -audiovisual- ya sea en escenarios IDES o IDMS), para la generación de contenido mulsemmedia, se deben realizar una serie de tareas adicionales para poder etiquetar los efectos sensoriales que deben ser presentados al usuario en momentos específicos del contenido multimedia. Para ello, se genera un fichero SEM a partir de la herramienta SEVino 2.0, descrita en [121]. Esta herramienta permite etiquetar, de una manera gráfica elementos y efectos sensoriales sobre cualquier contenido audiovisual y generar, de manera automática, el anteriormente mencionado fichero SEM. Sin embargo, tal y como se expone en el capítulo anterior, se asocia un valor PTS del contenido con un determinado efecto. Esto supone una problemática cuando el contenido audiovisual es transmitido a través de redes broadcast, ya que tal y como se ha expuesto, a lo largo de la cadena de transmisión valores como el PTS pueden verse modificados y por tanto, el SEM deja de tener coherencia con el contenido. Por tanto, una vez se ha generado el fichero con la herramienta SEVino 2.0, se modifica y añade una nueva propiedad para cada efecto etiquetado, denominada *global-timestamp*. Esta propiedad permite asociar a cada efecto un instante temporal global, basado en la misma línea temporal insertada en el descriptor TEMI *timeline* del contenido principal.

Una vez el fichero SEM está generado correctamente, se almacena en un servidor convencional en la parte del proveedor de contenidos de la plataforma y, a su vez, su ubicación es señalizada en el fichero RMCF con la etiqueta <MEDIA> y especificando que se trata de contenido mulsemmedia en la propiedad *media_type*="SEM". De esta manera, la plataforma ya está habilitada para proporcionar contenido mulsemmedia al consumidor o usuario final.

5.3.2 MÓDULO DE GENERACIÓN DEL FICHERO RELATED MEDIA CONTENT FILE

Este módulo es el encargado de generar el fichero RMCF, en el que se definen (metadatos) y señalizan todos los contenidos (complementarios y relacionados con el contenido broadcast o principal) disponibles a través de redes broadband, además de otro tipo de información necesaria (como, por ejemplo, la localización del reloj global), mediante la setiquetas definidas en la Tabla 2 del Capítulo 4.

5.3.3 MÓDULOS DE TRANSMISIÓN DE CONTENIDOS

A continuación se exponen los módulos encargados de la transmisión de los contenidos broadcast y broadband.

MÓDULO DE TRANSMISIÓN DE CONTENIDO BROADCAST

Una vez el contenido se ha generado en formato MPEG2-TS con la señalización necesaria integrada en los descriptores indicados, se procede a su transmisión mediante tecnología broadcast (DVB-T). Para ello se utiliza un PC con una tarjeta interna (PCI) que se encarga de realizar la modulación y transmisión a través de una salida RF a la que se conecta un cable coaxial. Concretamente, en la plataforma desarrollada se ha utilizado un ordenador con S.O: Windows 10, procesador IntelCore i7, 16GB de memoria RAM; con una tarjeta PCI marca Dektec, modelo DTA-2111, y el software Dektec StreamXpress entregado con la misma (Fig. 50).

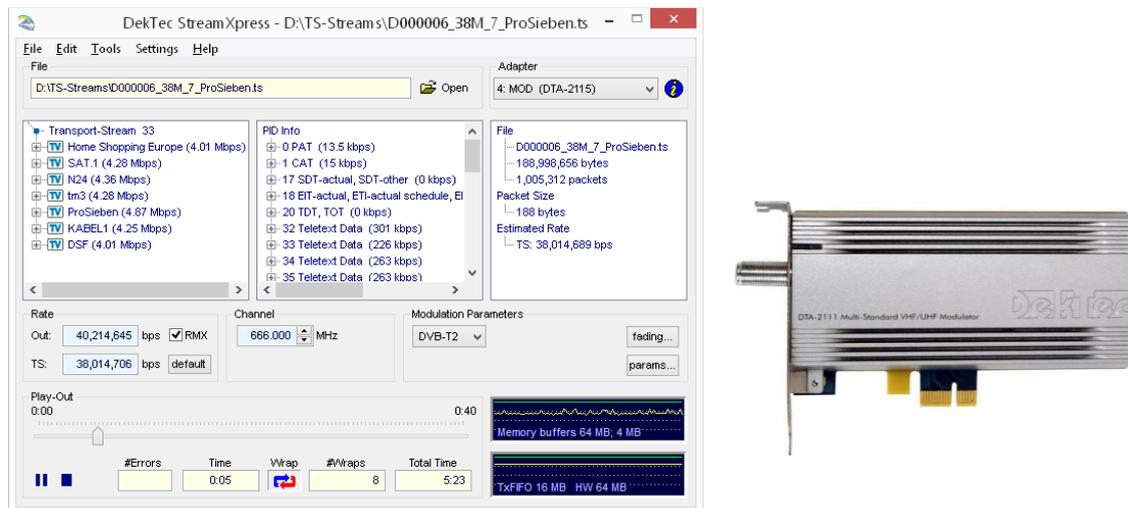


Fig. 50 Software StreamXpress y tarjeta moduladora DTA-2111, Dektec

5.3.3.1 MÓDULO DE TRANSMISIÓN DE CONTENIDO AUDIOVISUAL BROADBAND

El contenido audiovisual disponible a través de la red broadband se encuentra alojado en formato DASH en un servidor HTTP convencional encargado de proporcionar a los clientes conectados los segmentos que han solicitado en sus respectivas peticiones. La calidad del segmento solicitado (y por tanto, del contenido visualizado) corre a cargo del algoritmo de selección de calidad definido en el propio reproductor del cliente. En cuanto a la transmisión de contenido a través de la red broadband, no existe ninguna diferencia significativa entre que el contenido solicitado sea convencional u omnidireccional.

5.4 PARTE DEL CONSUMIDOR O USUARIO FINAL

En la parte del consumidor o usuario final (lado derecho de la Fig. 41), se puede observar los diferentes módulos y elementos involucrados. Todos estos elementos son necesarios para el descubrimiento, selección, recepción, procesado y reproducción sincronizada de los contenidos híbridos disponibles. Además, estos contenidos pueden reproducirse en todos en un único terminal (MS y CS apps integrados en el mismo) o en dispositivos secundarios (incluyendo sólo la CS app). A continuación, se describen los módulos involucrados y sus funcionalidades.

Los dispositivos de consumo multimedia pueden integrar tanto la aplicación MS como la CS. Esto significa que un único dispositivo puede incluir ambas aplicaciones siempre y cuando el rendimiento del dispositivo lo permita. Tanto la MS como la CS app implementan los requisitos y las funcionalidades definidas en el estándar HbbTV en relación con el uso de segundas pantallas.

Las funcionalidades y mecanismos necesarios para conseguir IDMS se han integrado en la MS app. De esta manera, el contenido principal (el que reproduce la MS app) es el que se utilizará como referencia a la hora de alcanzar un estado de sincronización IDMS satisfactorio. Esto permite que la sincronización IDES siga siendo posible, puesto que el contenido complementario disponible vía broadband se sincroniza respecto al contenido broadcast de la MS app y por tanto, el proceso de sincronización IDMS es transparente al mecanismo IDES.

Cuando en el escenario del caso de uso a implementar en la plataforma se requiera IDMS, la MS app extraerá del fichero RMCF la información necesaria para establecer una conexión con el IDMS Manager indicado en el mismo. Dicha conexión se hace a través de Websocket, puesto que es una tecnología incluida en el estándar HbbTV y, por tanto, los dispositivos que cumplan con dicho estándar serían también capaces de implementar este canal de comunicación.

Una vez el usuario pasa a estar dentro de una sesión IDMS (cuyo procedimiento y mensajes se han descrito en el capítulo anterior), puede comenzar a enviar mensajes *IDMS Report* (y recibir mensajes *IDMS Settings*, en caso de pertenecer a una sesión IDMS cuyo esquema de sincronización sea SMS) para alcanzar un estado de sincronización IDMS satisfactorio. Cuando se recibe un mensaje *IDMS Settings* (en el caso SMS), o bien haya recibido los suficientes *IDMS Reports* (en el caso DCS), la MS app calcula el nivel de asincronía IDMS. En el caso de superar el umbral de sincronización IDMS máximo (*session threshold*), se deberá realizar un ajuste en el proceso de reproducción del contenido principal. Al tratarse de un ajuste que se lleva a cabo en el proceso de reproducción de un contenido que se ha recibido vía broadcast, se ha configurado en el mismo un búfer capaz de almacenar aproximadamente el equivalente de 12s de contenido con un bitrate de 20Mbps. Estos valores se han adoptado según la recomendación propuesta en el estándar HbbTV, sección 13.5.3. Además, este tamaño de búfer es suficiente para compensar los retardos (y las diferencias de retardos) existentes en la actualidad, tal y como se analiza en [17], [19].

5.4.1 TERMINAL HÍBRIDO – MÓDULO MAIN SCREEN

El terminal híbrido se encuentra en el dispositivo principal, y en él está integrado el módulo MS, que es el que ejecuta la MS app, el cual es el encargado de recibir, procesar y presentar el contenido MPEG2-TS recibido vía broadcast (DVB-T). Para este reproductor, se ha utilizado el framework GStreamer [24], el cual permite configurar un reproductor para contenido recibido a través de DVB-T. La elección del framework GStreamer se debe, entre otros factores, a la modularidad que proporciona respecto a cada una de las etapas de un reproductor, es decir, se pueden añadir los elementos de recepción, demultiplexación, decodificación y renderizado de manera individual a la hora de implementar un reproductor multimedia. Al reproductor resultante de este modelo de implementación, consistente en “concatenar” los elementos uno tras otro, GStreamer lo denomina *pipeline* (tubería).

MÓDULO DE REPRODUCCIÓN DE CONTENIDOS MPEG2-TRANSPORT STREAM (RED BROADCAST)

El resto de los componentes GStreamer involucrados en un reproductor de contenido DVB-T funcional, capaz de extraer los descriptores TEMI timeline y location se muestran en la Fig. 51.

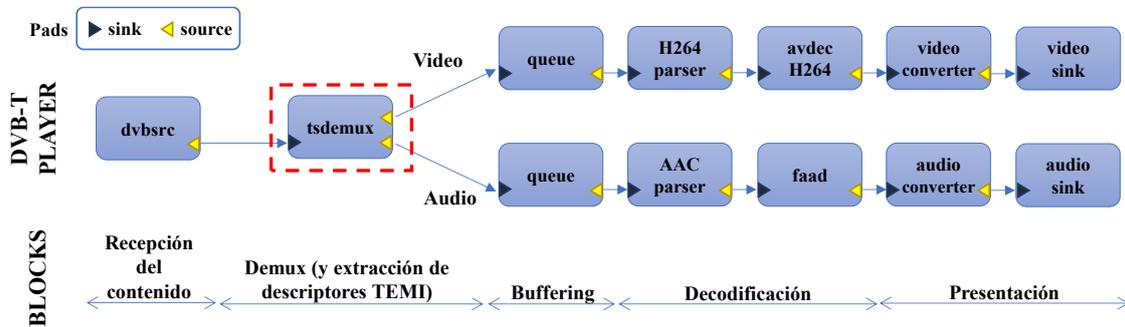


Fig. 51 Componentes GStreamer para la integración de un reproductor de contenido DVB-T

Tras el componente *tsdemux*, se obtienen, de manera separada, el audio y el vídeo. Para ambos casos, el proceso consiste en decodificar el contenido y presentar el contenido decodificado. El resto de los elementos se describen en la Tabla 11.

Tabla 11 Elementos GStreamer utilizados para la reproducción de contenido broadcast

Elemento (contenido)	Descripción	
<i>dvbsrc (DVB-T)</i>	Recibe, sintoniza y demodula la señal DVB-T recibida	
<i>tsdemux</i>	Procesa los flujos MPEG2-TS, identificando y separando (demultiplexando) los flujos y programas disponibles (audio, vídeo, texto, etc.)	
<i>queue</i>	Normalmente se utiliza como búfer o cola de datos, pero también se puede utilizar para estabilizar las salidas de un demultiplexor (por ejemplo, <i>tsdemux</i>) y mejorar el rendimiento general. Esta última es la funcionalidad por la que se ha integrado este elemento en los reproductores de la presente plataforma.	
Audio Decoder Block	<i>aacparse</i>	Analizador del flujo de audio codificado AAC
	<i>faad</i>	Decodificador AAC
	<i>audioconvert</i>	Convertor de audio (si resulta necesario)
Video Decoder Block	<i>h264parse</i>	Analizador de flujo de vídeo codificado H.264
	<i>avdec_h264</i>	Decodificador H.264
	<i>videoconvert</i>	Convertor de vídeo (si resulta necesario)
<i>autovideosink</i>	Selecciona automáticamente la salida de vídeo más apropiada según el dispositivo en uso, su hardware y su sistema operativo (por ejemplo, X11, GL, etc.)	
<i>autoaudiosink</i>	Selecciona automáticamente la salida de audio más apropiada según el dispositivo en uso, su hardware y su sistema operativo (por ejemplo, alsa, pulseaudio, etc.)	

MÓDULO DE RECEPCIÓN DE CONTENIDOS MULTIMEDIA RECIBIDOS A TRAVÉS DE LA RED BROADCAST

Como ejemplo, se muestra el código de desarrollo para configurar la recepción de contenido DVB-T, en la frecuencia de 770MHz (la utilizada actualmente por RTVE para transmitir sus canales en TDT en la región del sur de Valencia), empleando para ello el componente de GStreamer denominado *dvbsrc* (Fig. 52).

```
dvb.dvbsrc = gst_element_factory_make ("dvbsrc", "dvbsrc");
g_object_set (G_OBJECT (dvb.dvbsrc), "adapter", 0, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "modulation", 6, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "trans-mode", 2, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "bandwidth-hz", 8000000, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "code-rate-hp", 9, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "code-rate-lp", 9, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "guard", 4, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "hierarchy", 4, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "inversion", 2, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "frequency", 770000000, NULL);
dvb.dvbsrc = gst_element_factory_make ("dvbsrc", "dvbsrc");
g_object_set (G_OBJECT (dvb.dvbsrc), "adapter", 0, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "modulation", 6, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "trans-mode", 2, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "bandwidth-hz", 8000000, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "code-rate-hp", 9, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "code-rate-lp", 9, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "guard", 4, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "hierarchy", 4, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "inversion", 2, NULL);
g_object_set (G_OBJECT (dvb.dvbsrc), "frequency", 770000000, NULL);
```

Fig. 52 Código implementado para la configuración del receptor DVB-T.

MÓDULO DE EXTRACCIÓN Y ANÁLISIS DE LA SEÑALIZACIÓN DE CONTENIDO MULTIMEDIA Y MULSEMEDIA COMPLEMENTARIO AL CONTENIDO BROADCAST RELACIONADO Y COMUNICACIÓN A LOS COMPANION SCREEN Y CONTROLADOR DE EFECTOS

En el caso particular de la MS app, al insertar los descriptores TEMI, se ha modificado el código fuente del elemento del *pipeline* encargado del desencapsulado de los flujos MPEG2-TS (denominado *tsdemux*), ya que es en ese punto en el que se va a extraer la información correspondiente a los descriptores TEMI. En concreto, se ha añadido la capacidad de poder extraer (y analizar) ambos descriptores TEMI (*location* y *timeline*) en tiempo real, tan pronto como el elemento *tsdemux* del *pipeline* recibe el flujo MPEG2-TS. En la Fig. 53 se muestra el código relacionado con la monitorización de las marcas temporales insertadas en el contenido MPEG2-TS. En particular, se crea una llamada asíncrona (*callback*) que se dispara una vez se obtiene un nuevo valor del descriptor TEMI *timeline* en el bus nativo de GStreamer, utilizado para la transmisión de (meta)datos entre componentes de GStreamer.

```
dvb.tsdemux = gst_element_factory_make ("tsdemux", "tsdemux");
g_object_set (G_OBJECT (dvb.tsdemux), "emit-stats", TRUE, NULL);
dvb.bus = gst_element_get_bus (dvb.pipeline);
gst_bus_add_signal_watch (dvb.bus);
g_signal_connect (G_OBJECT (dvb.bus), "message::element", (GCallback)pts_cb, &dvb);
```

Fig. 53 Código implementado como callback para recibir la información insertada en el descriptor TEMI

El elemento de GStreamer encargado de desencapsular los flujos MPEG2-TS (*tsdemux*) ha sido modificado para poder extraer los descriptores TEMI (*location* y *timeline*) y de esta manera proporcionar la información necesaria para conocer qué tipos de contenidos adicionales están

disponible a través de la red broadband y poder sincronizar estos contenidos respecto al contenido principal presentado en la MS app. En la plataforma, el descriptor TEMI *location* almacena la URI del fichero RMCF, el cual está almacenado en un servidor web convencional en la parte del proveedor de contenidos de la plataforma. Para un escenario IDEs con contenido multimedia híbrido, en el fichero RMCF debe haber (al menos), una etiqueta <CLOCK> que indique la ubicación del servidor de reloj global (NTP) utilizado y una etiqueta <MEDIA> que señale la ubicación de contenido multimedia adicional disponible vía broadband. El procedimiento de cómo y cuándo la MS app intercambia esta información con la CS app se puede consultar en el Capítulo 4.

Respecto al contenido mulsemmedia, en el fichero RMFC debe estar indicada la ubicación del fichero SEM, el cual almacena las etiquetas de efectos multisensoriales existentes durante la reproducción del contenido definido. Se deberá analizar la propiedad *media_type* de la etiqueta <MEDIA> dentro del fichero RMFC, cuando su valor es igual a "SEM". Una vez localizada su ubicación, se consulta el fichero SEM y se analiza siguiendo su contenido, tal y como se define en el Capítulo 4.

Una vez las MS y CS apps han establecido correctamente el canal de comunicación bidireccional a través del WSS App2App del terminal híbrido de HbbTV, ubicado en el mismo dispositivo que la MS app, la CS app recibe por parte de la MS app la información necesaria para poder consumir un contenido audiovisual complementario disponible vía broadband. En concreto, recibe información referente al reloj global que debe ser utilizado y la ubicación (URL) del (los) contenido(s) disponible(s).

MÓDULO DE EMULACIÓN DE UN TERMINAL HÍBRIDO COMPATIBLE CON HYBRID BROADCAST BROADBAND TELEVISION 2.0

Durante el desarrollo de esta Tesis Doctoral, no se ha podido contar con un TV completamente compatible con el estándar HbbTV en la versión 2.0, que cuenta con funcionalidades de interacción con segundas pantallas. Es por este motivo, que se ha hecho uso de la herramienta proporcionada por Fraunhofer denominada *Node.js hbbtv module*¹². Para ello, ha sido necesario instalar el framework de Node.js para poder ejecutar dicha herramienta. Este módulo incluye un servidor DIAL para el descubrimiento del terminal híbrido por parte de las CS apps y, por otro lado, dispone de los canales de comunicación bidireccionales basados en Websockets denominados *App2App* en la terminología del estándar HbbTV. A través de este canal de comunicación se transmitirá la información acerca de los contenidos complementarios disponibles y la información temporal a la que las CS apps deben sincronizarse.

5.4.2 MÓDULO COMPANION SCREEN

A continuación, se describe el módulo CS, que es el encargado de ejecutar la CS app, que incluye diferentes procesos explicados en diferentes subapartados.

PROCESO DE DESCUBRIMIENTO DE DISPOSITIVOS O SERVICIO Y LANZAMIENTO DE LA APLICACIÓN SECUNDARIA

La plataforma integra las funcionalidades de descubrimiento y vinculación entre los dispositivos involucrados, de manera transparente al consumidor. Para ello, el descubrimiento, asociación e interacción entre la MS app y una CS app y/o un controlador de efectos mulsemmedia debe realizarse de manera automática y sin que el usuario forme parte de este proceso con el fin de proporcionar una experiencia de usuario lo más satisfactoria posible.

¹² <https://github.com/fraunhoferfokus/node-hbbtv>

En particular, se ha integrado un servidor DIAL en el dispositivo ejecutando la MS app, contenido en la herramienta de emulación del terminal híbrido compatible con HbbTV2.0. Dicho servidor está continuamente esperando peticiones de clientes DIAL (CS apps y/o controlador de efectos multimedia). Por otra parte, las funcionalidades correspondientes al protocolo SSDP han sido implementadas desde cero en las CS apps. En concreto, con el fin de descubrir algún servicio o dispositivo DIAL, la CS app envía un mensaje del tipo M-SEARCH (del protocolo SSDP) a la dirección multicast 239.255.255.250:1900, en la cual está el MS (el servidor DIAL) escuchando posibles peticiones de descubrimiento. Tras escuchar una petición de descubrimiento, la MS app responde de manera unicast a la CS app que ha realizado la petición (línea continua entre MS y CS apps en la Fig. 16), proporcionando la información necesaria para que ambos puedan establecer un canal de comunicación bidireccional (línea discontinua entre MS y CS apps en la Fig. 16) basado en la tecnología WS, a través del servidor (WSS, del inglés WebSocket Server) App2App. Este canal de comunicación estará activo durante toda la sesión de consumo multimedia. En la Fig. 54 se muestra, como ejemplo, el código desarrollado para realizar una petición de descubrimiento en dispositivos Android (lenguaje Java).

```
/**
 * Clase encargada de crear un socket para enviar y recibir mensajes multicast al emplear el protocolo SSDP
 */
public class SSDPSocket {
    /** Socket multicast */
    MulticastSocket mSSDPSocket;
    /** Direccion multicast empleada en el mensaje M-SEARCH de SSDP*/
    InetAddress broadcastAddress;

    /** Constructor de la clase, inicializa el socket multicast y lo configura */
    public SSDPSocket() throws IOException {
        mSSDPSocket = new MulticastSocket();
        broadcastAddress = InetAddress.getByName(SSDPConstants.ADDRESS);
        mSSDPSocket.joinGroup(broadcastAddress);
        Log.d("SSDPSocket", "SSDPSocket creado");
    }

    /** Metodo que permite el envio de mensajes multicast - SSDP */
    public void send(String data) throws IOException {
        DatagramPacket dp = new DatagramPacket(data.getBytes(), data.length(),
            broadcastAddress,SSDPConstants.PORT);

        mSSDPSocket.send(dp);
        Log.d("SSDPSocket", "Mensaje recien enviado");
    }

    /** Metodo que permite la recepcion de mensajes multicast - SSDP */
    public DatagramPacket receive() throws IOException {
        byte[] buf = new byte[1024];
        DatagramPacket dp = new DatagramPacket(buf, buf.length);
        Log.d("SSDPSocket", "Esperando paquete respuesta...");
        mSSDPSocket.receive(dp);

        return dp;
    }
}
```


	queue	Normalmente se utiliza como búfer o cola de datos, pero también se puede utilizar para estabilizar las salidas de un demultiplexor (por ejemplo, <i>tsdemux</i>) y mejorar el rendimiento general. Esta última es la funcionalidad por la que se ha integrado este elemento en los reproductores de la plataforma desarrollada.
Audio Decoder Block	<i>aacparse</i>	Analizador del flujo de audio codificado AAC
	<i>faad</i>	Decodificador AAC
	<i>audioconvert</i>	Convertor de audio (si resulta necesario)
Video Decoder Block	<i>h264parse</i>	Analizador de flujo de vídeo codificado H.264
	<i>avdec_h264</i>	Decodificador H.264
	<i>videoconvert</i>	Convertor de vídeo (si resulta necesario)
	autovideosink	Selecciona automáticamente la salida de vídeo más apropiada según el dispositivo en uso, su hardware y su sistema operativo (por ejemplo, X11, GL, etc.)
	autoaudiosink	Selecciona automáticamente la salida de audio más apropiada según el dispositivo en uso, su hardware y su sistema operativo (por ejemplo, alsa, pulseaudio, etc.)

MÓDULO REPRODUCTOR DE CONTENIDO OMNIDIRECCIONAL

En la parte del consumidor o usuario final, se ha diseñado un reproductor basado en web específico y compatible con vídeo omnidireccional que, además, permite reproducir contenido tradicional 2D. Dicho reproductor está diseñado para poder ser ejecutado en dispositivos secundarios HMD y aprovechar la interactividad e inmersividad que estos ofrecen. Sin embargo, también se puede ejecutar en dispositivos más comunes (o tradicionales) como tablets, smartphones u ordenadores, en los que el contenido se visualiza en una pantalla plana y para desplazar el punto de vista, se debe utilizar los sensores integrados en los propios dispositivos (giroscopio y acelerómetros) o el dedo o el ratón del usuario.

El diseño del reproductor basado en web consta de dos partes. Por un lado, se incluye un salón virtual que puede verse en sus 360 grados, con el fin de proporcionar al usuario una experiencia lo más inmersiva posible, con 3 grados de libertad (3DoF). En este salón virtual el contenido del vídeo que se puede visualizar es 2D y se reproduce en una TV virtual. Por otro lado, también incluye un reproductor omnidireccional, el cual es independiente del salón virtual y proporciona al consumidor o usuario final una reproducción en 360° del contenido omnidireccional. El proceso de cambio entre ambas partes se lleva a cabo a través de botones o puntos de interacción que han sido implementados tanto en el salón virtual como en el reproductor omnidireccional.

Se ha recreado un escenario lo más inmersivo posible. Se ha diseñado un escenario que emula un salón virtual con una TV, donde se podrá consumir el contenido tradicional 2D como en un hogar convencional. Por otro lado, se puede consumir el contenido omnidireccional al seleccionarlo, en el caso de que esté disponible y señalado en dicho escenario en la parte derecha del mismo y con un indicativo de 360°. Como este escenario está diseñado para poder ser reproducido con HMD, cuenta con diferentes puntos de interacción, con el fin de cambiar de vista o de contenido, ya sea 2D u omnidireccional. La Fig. 56 muestra, para la visualización de contenido omnidireccional, las regiones del espacio donde se ubicará el contenido audiovisual (en gris claro), tanto en el salón virtual (región con la TV virtual) como en la reproducción de contenido 360° propiamente dicho.

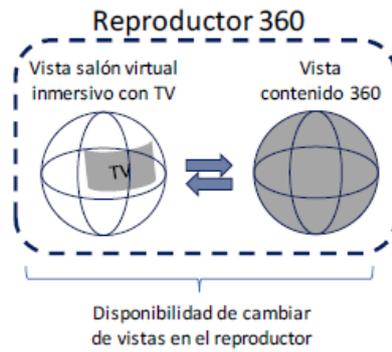


Fig. 56 Ubicación del contenido audiovisual (gris claro) en el salón virtual (izquierda) y en el reproductor omnidireccional (derecha)

En la Fig. 57 se puede observar el conjunto de funcionalidades y puntos de interacción que ofrece el escenario inmersivo del salón virtual. El aspecto del escenario inmersivo del salón virtual y del reproductor omnidireccional puede verse en las Fig. 57 y Fig. 58, respectivamente.



Fig. 57 Aspecto del escenario inmersivo del salón virtual visualizado en una pantalla tradicional (a) y en un HMD (b)

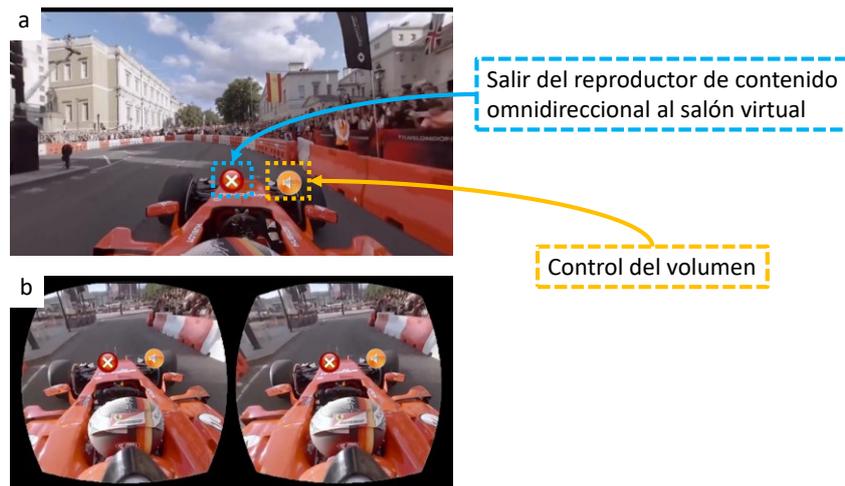


Fig. 58 Aspecto final del reproductor omnidireccional en una pantalla tradicional (a) y en un HMD (b)

Como se puede observar en la Fig. 57, el escenario inmersivo que emula el salón cuenta con una pantalla virtual de TV, a través de la cual se puede consumir contenido tradicional 2D. Si existe contenido omnidireccional disponible, el usuario puede seleccionarlo y reproducirlo, cambiando así del escenario inmersivo al reproductor omnidireccional propiamente dicho. Dependiendo del dispositivo utilizado, el contenido omnidireccional se puede visualizar de un modo u otro: de una forma tradicional (en pantalla plana), o bien de forma estereoscópica (en HMD). La forma en la que el consumidor o usuario final interactúa con el escenario o con el reproductor omnidireccional, dependerá del dispositivo secundario que utilice como, por ejemplo, con teclado y ratón (si es un ordenador), con el dedo (si es una pantalla táctil) o a partir de colisiones con puntos de interés (en caso de utilizar un HMD). La Fig. 59 muestra un diagrama de flujo que define el comportamiento del reproductor omnidireccional en función del dispositivo que la ejecuta en el instante de cargar la GUI (Graphical User Interface - Interfaz Gráfica de Usuario) correspondiente.

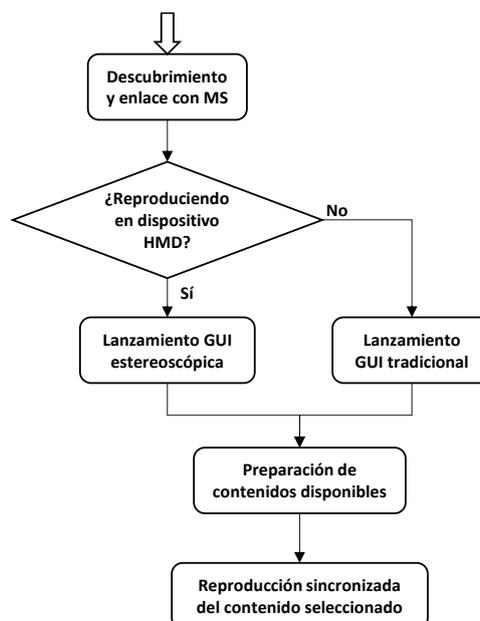


Fig. 59 Selección del GUI del reproductor omnidireccional en función del dispositivo que lo ejecuta

Para llevar a cabo el desarrollo de este escenario inmersivo para soportar contenido omnidireccional, se han adoptado tecnologías web, puesto que, tal y como se expone en [158],

[159], el uso de este tipo de tecnologías permiten unificar una sola aplicación para múltiples dispositivos con sus respectivas especificaciones. Además, tal y como se expone en [160], el desarrollo de aplicaciones con tecnologías web reduce los tiempos de desarrollo, mantenimiento y distribución. Además, por un lado, garantiza un soporte multidispositivo, multiplataforma y multinavegador y, por otro lado, es menos sensible a restricciones asociadas a determinadas políticas de bloqueo de firewalls o problemas relacionados con la traducción de direcciones de red (Network Address Translation o NAT).

De hecho, en [158] se propone el uso de este tipo de tecnologías para proporcionar a broadcasters y generadores de contenido para la TV Híbrida una solución de desarrollo unificado, abstrayendo así el tipo de dispositivo para el que se va a desarrollar la aplicación híbrida. En [159], se propone una metodología para generar interfaces de usuario para aplicaciones en escenarios multipantalla y multidispositivo.

Por tanto, este escenario inmersivo es compatible con la gran mayoría de dispositivos de consumo modernos, que además puedan ser integrables en dispositivos HMD o bien utilizables como una pantalla plana común. En particular, la implementación se ha llevado a cabo utilizando, principalmente, las librerías JavaScript *dash.js* [161] y *three.js* [162], Fig. 60.

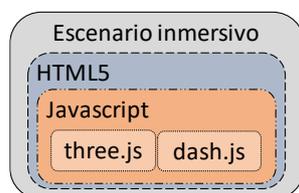


Fig. 60 Tecnologías utilizadas en el desarrollo del escenario inmersivo

Por un lado, la librería *dash.js* permite la reproducción de contenido basado en DASH en navegadores. Por otro lado, la librería *three.js* permite la generación y gestión gráfica en los navegadores, por lo que se puede combinar con otros elementos HTML (por ejemplo, el *canvas*). Así, se ha podido implementar el escenario virtual del salón y se ha podido proyectar el contenido omnidireccional en una vista de 360 grados. Adicionalmente, el uso de esta librería también permite añadir una presentación estereoscópica tanto en el salón virtual como en el reproductor omnidireccional cuando se utiliza un dispositivo HMD. De hecho, permite desplazar una de las vistas estereoscópicas para que cada ojo experimente la misma escena con cierta distancia, con el fin de proporcionar una sensación de profundidad.

En la implementación del salón virtual, la librería *three.js* se ha utilizado para gestionar una malla geométrica esférica, donde se proyecta la textura del salón omnidireccional. Para poder proporcionar al usuario 3 grados de libertad (3DoF), se sitúa la cámara (o punto de visión del usuario) en el centro del escenario. Además, se sitúa una malla geométrica rectangular, que adoptará el papel de emular la TV para reproducir el contenido 2D. Para ello, se proyectará en ella el contenido 2D haciendo uso de la librería *dash.js*.

Para interactuar con el escenario virtual y las notificaciones (señalizaciones) de otros contenidos disponibles, se han generado capturas (imágenes) de dichos contenidos disponibles, que se muestran como iconos colocados verticalmente a la derecha de la TV virtual. Se ha añadido la funcionalidad de interactuar con dichos iconos a través de interacciones clásicas (como el ratón o pantalla táctil) o a través de colisiones (por ejemplo, utilizando *raycasting* -monitorizando el impacto del FoV del usuario con la superficie de un icono determinado-), para que se disparen los eventos asociados, como cambiar de vista, de tipo de reproductor o quitar el sonido (mute).

Para reproducir contenido omnidireccional, también se ha hecho uso de la librería *three.js*. Dependiendo del tipo de proyección del contenido (ERP o CMP), se genera una malla geométrica

rectangular o esférica, respectivamente. En ambos casos, la textura que cubre esta malla es el contenido omnidireccional, que se reproduce gracias a la librería *dash.js*.

MÓDULOS RELACIONADOS CON EL CONTROL DE EFECTOS MULTISENSORIALES

Los módulos integrados en la parte del consumidor o usuario final están directamente relacionados con el tipo de efectos que van a estar disponibles durante la experiencia de consumo.

En este apartado se presentant todos los módulos de la plataforma relacionados con la gestión y el control de la generación o activación de los efectos multisensoriales. Cabe resaltar el papel del Mulsemmedia Controller (MC), previamente definido en el Capítulo 4 y encargado de interactuar con los dispositivos generadores de efectos sensoriales una vez recibe una petición determinada por parte del MS o CS app.

El procedimiento para presentar contenido mulsemmedia depende de la MS o CS app, por lo que el MC y la MS o CS app deben haberse descubierto previamente. Esto se lleva a cabo a través del servicio de descubrimiento definido en DIAL, de manera similar al procedimiento que realiza una CS app para descubrir a la MS app. Sin embargo, una vez la MS o la CS app y el MC conocen la dirección IP (local) del otro, no establecen un canal de comunicación bidireccional a través del servidor App2App, sino que la MS o la CS app se comunica con el MC a través de una API REST, tal y como se ha descrito en el Capítulo 4.

En la plataforma el MC se ha implementado en un dispositivo SoC (System on Chip, por ejemplo, un dispositivo Arduino) y su única función es la de enlace entre la MS o CS app y el dispositivo genrador de cada efecto multisensorial, sin necesidad de realizar ningún tipo de procesado más que el de activar o desactivar determinados pines de salida, a los que estarán conectados relés para activar o desactivar determinados efectos. En la plataforma, los efectos que se han incluido son luces, viento, vibración y aromas. Como MC, se ha utilizado un microcontrolador NodeMCU¹³, basado en el SoC Esp8266, que cuenta con conectividad WiFi integrada (Fig. 61).



Fig. 61 Microcontrolador NodeMCU

Adicionalmente, este microcontrolador puede implementar proyectos a través del lenguaje de programación en el que se basa el entorno de desarrollo de Arduino, denominado *Processing*¹⁴. Dicho lenguaje se compone de dos bloques principales, uno (incluido en la función denominada *setup()*) que se ejecuta únicamente al inicio de la ejecución del programa y otro que se ejecuta de manera continua y en bucle (incluido en la función denominada *loop()*).

Una vez la MS o la CS app ha recibido y analizado los efectos mulsemmedia etiquetados en el fichero SEM señalado en el RMCF, almacena el instante de generación en el que deben dispararse y, en dicho instante, realiza una petición REST al MC, activando o desactivando el efecto correspondiente.

¹³ https://www.nodemcu.com/index_en.html

¹⁴ <https://processing.org/>

Al tratarse el dispositivo adoptado para la función de MC, de un microcontrolador, debe evitarse que la mayor parte del procesado computacional sea llevado a cabo por esta entidad. Por lo tanto, el MC sólo realiza dos tareas: i) proporcionar una API REST, tal y como se ha explicado; y ii) emplear sus pines de conexión para emitir señales digitales (1 o 0), o analógicas (0 a 255). En estos pines es donde se conectan los dispositivos físicos encargados de generar los diferentes efectos multimedia (por ejemplo, activar un ventilador para emular viento, un emisor de aromas o un subwoofer de baja frecuencia para hacer vibrar el sillón o sofá donde esté el usuario). La Fig. 62 muestra las principales acciones que lleva a cabo el MC en un diagrama de flujo.

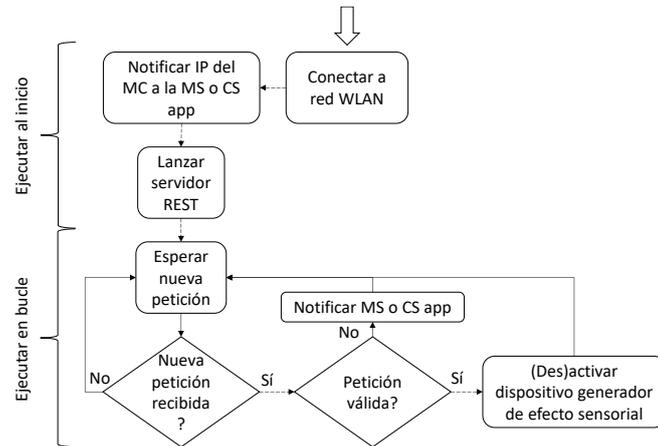


Fig. 62 Diagrama de flujo que muestra el comportamiento del MC

Cada efecto está asociado a un dispositivo diferente, por ejemplo, el efecto de viento a un ventilador de uso común, y un aroma a un ventilador con un dispensador de un olor determinado. Dichos dispositivos pueden alimentarse a través de una fuente de tensión continua o a la red eléctrica convencional. A continuación, se presentan los circuitos involucrados en la gestión y presentación de los diferentes efectos multisensoriales considerados en la implementación de la plataforma.

MÓDULO DE GENERACIÓN DE EFECTOS DE VIENTO Y DE CALOR

Para emular los efectos de viento y calor, se ha implementado el circuito de la Fig. 63, puesto que se ha utilizado un ventilador o un calefactor (según el efecto viento o calor, respectivamente) conectado a la red eléctrica y activados a través del relé normalmente abierto.

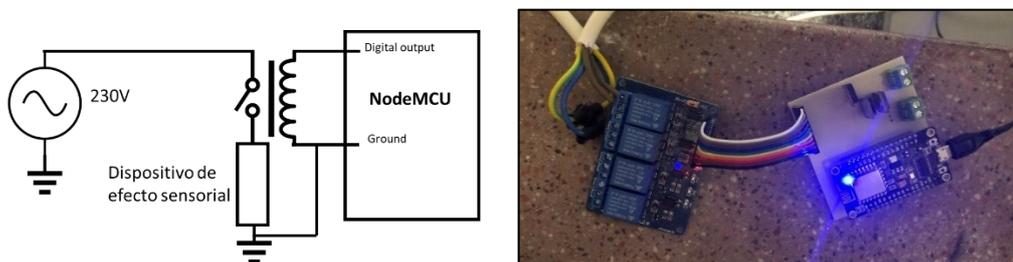


Fig. 63 Esquema (izquierda) y vista parcial del circuito electrónico (derecha) encargado de actuar sobre los dispositivos encargados de generar los efectos sensoriales de viento o calor.

Este circuito cuenta con un relé normalmente abierto que se acciona a partir de una salida digital del MC. Una vez se activa el relé, se cierra el circuito que alimenta el dispositivo de efecto sensorial (viento o calor) y, por tanto, este queda activo hasta que se desactive el relé a partir de dicha salida digital.

MÓDULO DE GENERACIÓN DE AROMAS

Con el objetivo de disparar o emitir aromas, se ha utilizado una serie de pequeños contenedores de partículas aromáticas, de manera que al situar un ventilador DC24V detrás y una carcasa envolvente adecuada y con alguna de las caras con rejillas, es capaz de transmitir el aroma hasta el consumidor. Está basado en el uso de los contenedores de aromas de la empresa Exhalia (actualmente Inhalio¹⁵). Los dispositivos electrónicos involucrados se han implementado de manera independiente, para poder contar con un mayor control sobre la gestión de este efecto, ya que el dispositivo generador de aromas proporcionado por Exhalia únicamente cuenta con conectividad USB, sólo es compatible con ciertos navegadores web (por ejemplo, Internet Explorer) y la distancia de emisión es corta (ventiladores poco potentes).

Debido a que, por la naturaleza de este efecto, con el objetivo de que el consumidor final aprecie, en mayor o menor grado, un determinado aroma, la intensidad del mismo debe ser un parámetro configurable. Por tanto, ha sido necesario hacer uso de la salida analógica del MC (esto es, salida PWM) para poder regular la intensidad del aroma (a partir de la velocidad de giro del ventilador).

Como el MC basado en NodeMCU puede transmitir hasta 5V de señal y esto no es suficiente para hacer funcionar correctamente los ventiladores de 24V DC, se han integrado unos controladores electrónicos de motores DC, como el presentado en la Fig. 64.

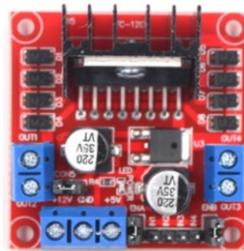


Fig. 64 Controlador PWM de motores DC

A partir de este dispositivo, se pueden controlar señales PWM de hasta 24V a partir de la salida PWM del MC, quedando un circuito como el que se muestra en la Fig. 65.

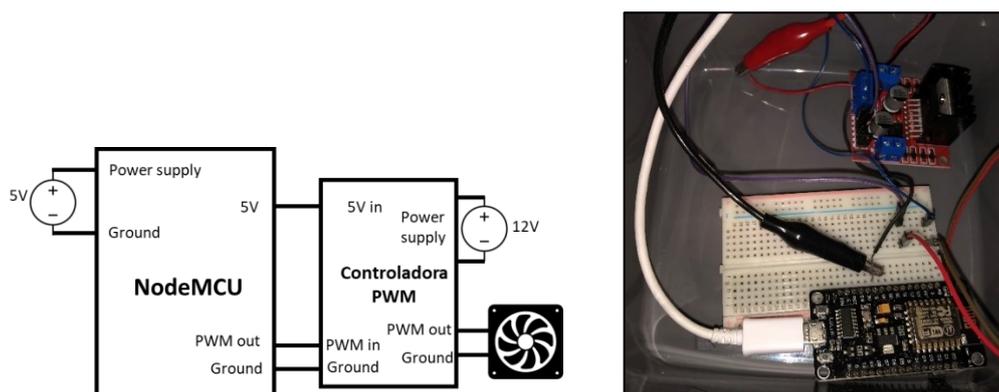


Fig. 65 Esquema (izquierda) y vista parcial del circuito electrónico (derecha) encargado de activar los ventiladores del dispositivo generador aromas

15

https://www.prweb.com/releases/digital_scent_leaders_inhalio_and_exhalia_merge_to_meet_growing_demand_from_mobility_rideshare_connected_home_hospitality_and_retail_markets/prweb16860030.htm

MÓDULO DE GENERACIÓN DEL EFECTO VIBRACIÓN

Para la implementación del efecto de vibración, se ha utilizado un altavoz de baja frecuencia o subwoofer (300Hz – 3KHz) Moncor BR-50, encargado de convertir la corriente eléctrica recibida en vibración. En este caso, ha sido necesario implementar un circuito que cuenta con un filtro paso bajo (Fig. 66) y un amplificador de la señal filtrada (Fig. 67). Cabe decir que ambas imágenes han sido obtenidas del proveedor, por lo que no se ha diseñado esta circuitería específicamente para la plataforma.



Fig. 66 Aspecto del filtro paso bajo

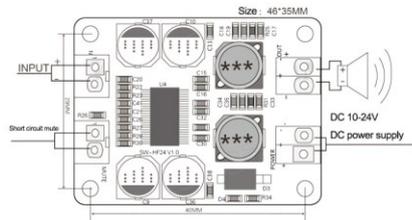


Fig. 67 Croquis del circuito amplificador de la señal

El circuito resultante se muestra en la Fig. 68. En dicha figura, las proporciones no son reales, sino que se ha presentado de esta manera para facilitar su comprensión.



Fig. 68 Esquema (izquierda) y vista parcial del circuito electrónico (derecha) encargado de la generación de vibraciones a partir de un subwoofer de baja frecuencia

MÓDULO GENERADOR DEL EFECTO LUZ AMBIENTAL

Para implementar la luz ambiental durante el consumo de contenido, se ha adoptado la herramienta Hyperion, disponible en dispositivos de bajo coste como la Raspberry Pi 3, que se basa en dividir en diferentes regiones los bordes o contornos del contenido y obtener su histograma (Fig. 69). De esta forma, se obtiene el color promedio para cada región, lo que permite enviar esta información a luces LED ubicadas a lo largo y ancho de la pantalla o TV, ampliando de esta manera el alcance visual del contenido que se está presentando. Se debe tener en cuenta que el array de luces LED a utilizar debe permitir la configuración independiente de cada una de las luces LED.

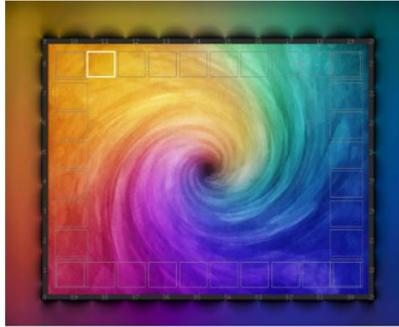


Fig. 69 Ejemplo de división del borde del contenido en regiones para obtener el valor de color promedio

Para implementar este efecto, se ha adquirido una serie de elementos que han permitido el tratamiento de la imagen a través de una Raspberry Pi 3 (Fig. 70):

- Splitter HDMI 1x2 (1 entrada, 2 salidas): permite conectar una salida directamente a la pantalla o TV, mientras que la otra salida se emplea para procesar la imagen y poder enviar la información necesaria al array de luces LED.
- Conversores HDMI a VGA y VGA a USB: la Raspberry Pi no tiene entrada VGA y la imagen debe ser recibida a través del puerto de entrada USB para poder ser procesada.
- Módulo Hyperion¹⁶: se encarga de analizar la imagen, dividir los bordes del contenido en regiones y obtener los histogramas, para calcular el promedio de color en cada una de las regiones. Se trata de una librería *open source* disponible en el enlace a pie de página que se usa comúnmente en dispositivos como la Raspberry Pi para crear luces ambientales alrededor de una pantalla o TV.
- Salidas GPIO de la Raspberry Pi: se encargan de comunicar al array de luces LED qué valor RGB deben emitir, extendiendo los colores del contenido a través de la luz ambiental generada.
- Array de luces LED con chipset WS2812: array de luces LED con un chipset específico que permite configurar de manera independiente cada una de las luces LED, permitiendo de esta forma la generación de la luz ambiental coherente con los colores del contenido que está siendo presentado.

¹⁶ <https://github.com/Hyperion-project>

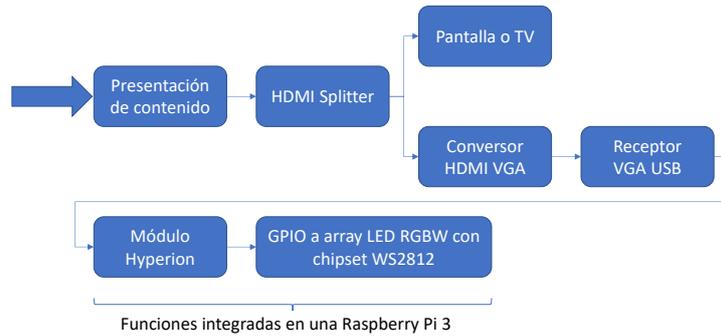


Fig. 70 Proceso de generación (arriba) y vista parcial del circuito implementado para la emisión de la luz ambiental coherente con el contenido presentado en el Main Screen

5.5 CONCLUSIONES

En este Capítulo se ha presentado la implementación de la plataforma extremo-a-extremo de generación y consumo de contenidos híbridos, multidestinataro, multidispositivo y multisensorial, en la que se han incluido todas las soluciones propuestas en la Tesis Doctoral

Se ha detallado el desarrollo su arquitectura y los módulos incluidos en la misma. También se ha descrito cada una de las funcionalidades incluidas en la misma: la generación del contenido multimedia, la inserción de los descriptores TEMI location y timestamp, la generación de etiquetas mulsemedia, la señalización de todo el contenido complementario disponible a través de la red broadband, así como toda la parte correspondiente a la recepción de los contenidos y la interacción entre los dispositivos de consumo en el extremo del usuario final.

Se han integrado en la plataforma los mecanismos de sincronización presentados en esta Tesis Doctoral, con el fin de evaluarlos, tanto objetiva como subjetivamente, en diferentes casos de uso, tal y como se expone en los dos capítulos siguientes.

Capítulo 6. EVALUACIÓN OBJETIVA Y RESULTADOS

En este capítulo se exponen la metodología y los resultados de las evaluaciones objetivas llevadas a cabo durante la Tesis Doctoral, con respecto al nivel de sincronización híbrida IDES e IDMS alcanzado en los escenarios de evaluación.

6.1 INTRODUCCIÓN

En la siguiente sección se presenta la evaluación objetiva de las soluciones de sincronización propuestas en la Tesis Doctoral en los siguientes escenarios:

- Un escenario básico incluyendo una solución IDES (denominado *escenario IDES básico*), el cual incluye contenido multivista;
- Un escenario inmersivo incluyendo una solución IDES (denominado *escenario IDES inmersivo*), el cual incluye contenido omnidireccional; y finalmente,
- Un escenario incluyendo soluciones de sincronización híbrida IDES e IDMS (denominado *escenario IDMS*).

Para todos ellos, primero se describe el escenario y, a continuación, la metodología seguida y los resultados obtenidos

6.2 ESCENARIO INTER-DISPOSITIVO BÁSICO

Para llevar a cabo la evaluación objetiva de las funcionalidades IDES de la plataforma se ha implementado uno de los escenarios con más popularidad [6] en cuanto al consumo de contenidos relacionados en diferentes pantallas y de forma sincronizada: un escenario multipantalla (multidispositivo) y multivista. En el mismo, el contenido de TV principal es transmitido a través de la red broadcast (DVB-T) y, adicionalmente, se dispone de cuatro contenidos complementarios relacionados con el mismo, accesibles a través de una red broadband (IP). La Fig. 71 muestra una visión general del escenario IDES básico (en el que no se incluyen dispositivos HMD para ejecutar la CS app).

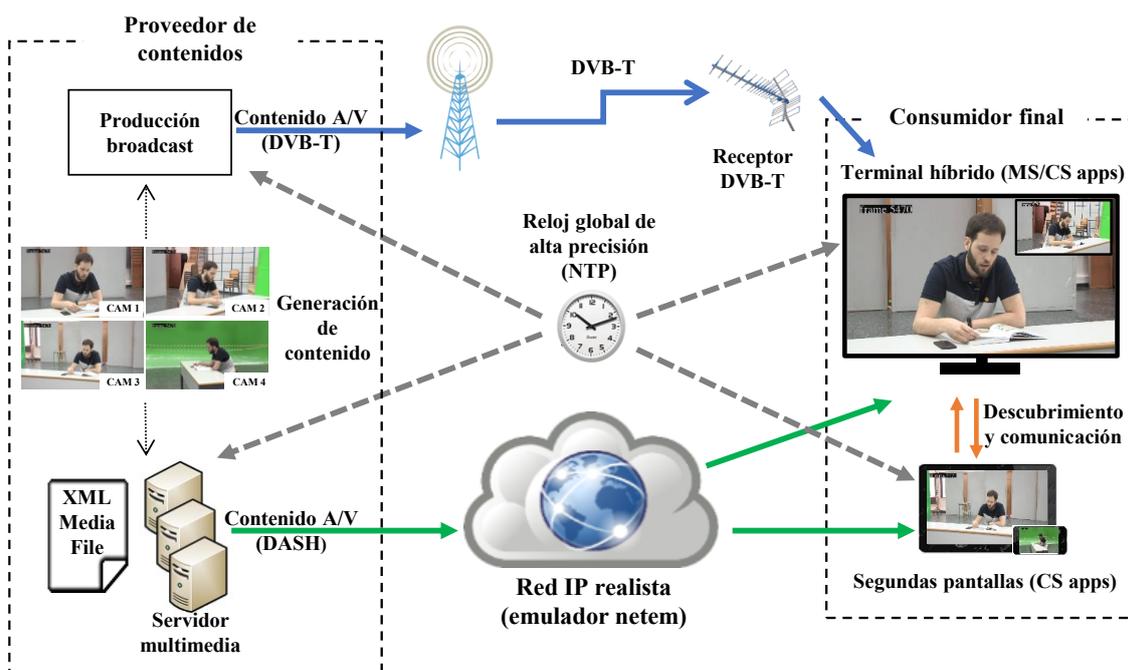


Fig. 71 Escenario inter-dispositivo implementado en la plataforma para la evaluación objetiva de la solución IDES básica (sin incluir HMD)

6.2.1 ESCENARIO DE EVALUACIÓN

Como se ha indicado, en este escenario, en la plataforma se ha implementado un caso de uso multivista y multidispositivo, el cual tiene el potencial de ofrecer al usuario una experiencia de consumo altamente personalizable e inmersiva [6]. Para esta evaluación, se ha creado una serie de contenidos multimedia cuyo tema ha sido la lectura de un libro. Estos contenidos han sido generados en las instalaciones del Campus de Gandia de la UPV, donde se dispone de la infraestructura y medios necesarios para ello. En concreto, se ha grabado la misma escena desde cuatro ángulos diferentes. Esto permite, por ejemplo, que se pueda transmitir una de las vistas a través de la red broadcast (como contenido principal) y, además, al mismo tiempo, disponer de las cuatro vistas a través de la red broadband. Para esta evaluación objetiva, las cuatro vistas se han multiplexado en el mismo flujo MPEG2-TS como cuatro flujos elementales (*Elementary Streams*, -ES-) y se han transmitido vía DVB-T a través de un canal Ultra High Frequency (UHF). La transmisión DVB-T se ha configurado con parámetros típicos: transmisión de canal UHF a 770MHz, mapeado 64-QAM, ancho de banda 8MHz, 1/32 de intervalo de guarda, modulación OFDM 8K y tasa de codificación 7/8).

Este flujo de transporte MPEG2-TS es recibido y demultiplexado por el receptor DVB, y una de las cuatro vistas disponibles puede ser seleccionada para ser reproducida por la MS app. Por otro lado, el contenido generado de las cuatro vistas también se ha codificado y empaquetado como contenido MPEG-DASH en un servidor web, para poder ser descargado de forma adaptativa a través de la red broadband, con el objetivo de ser reproducido por las CS apps. Por tanto, tal y como se ha expuesto en capítulos anteriores, por un lado, el contenido principal a ser reproducido en la MS app se codifica, encapsula y transmite vía DVB-T, mientras que, por otro lado, los contenidos relacionados complementarios se preparan y almacenan en un servidor web convencional para estar disponibles como contenido DASH.

Tal y como se ha indicado en los Capítulo 4 y Capítulo 5, un dispositivo o terminal híbrido puede reproducir tanto el contenido principal (MS app), como los contenidos secundarios (CS app) simultáneamente. Cuando se reproduce el contenido en el terminal híbrido, el (o los) contenido(s) se pueden presentar en forma de mosaico, a pantalla completa o en modo Picture-in-Picture (PiP),

Fig. 72. Adicionalmente, se ha añadido un menú en la aplicación desarrollada que permite seleccionar, dinámicamente, de entre los distintos contenidos disponibles (en este caso, diferentes ángulos o puntos de vista del mismo evento), Fig. 73.



Fig. 72 Aspecto de las aplicaciones MS y CS desarrolladas para la plataforma, tanto en el terminal híbrido como en el dispositivo secundario

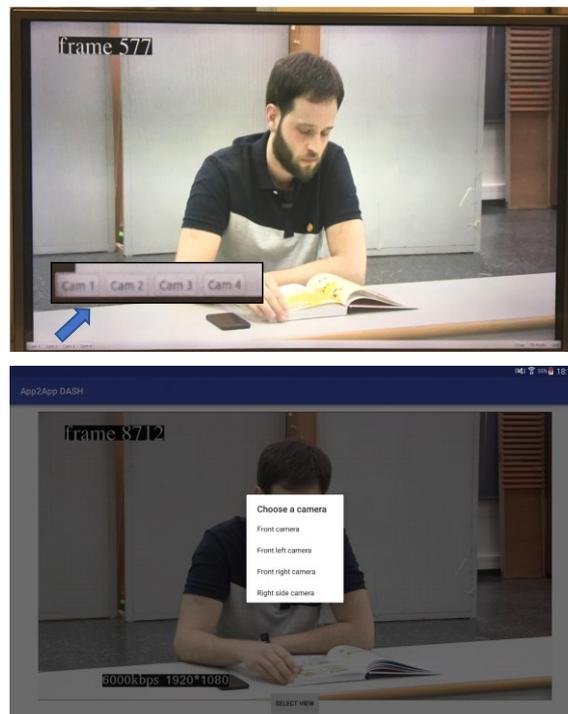


Fig. 73 Menú emergente para seleccionar el contenido broadband disponible en MS app (superior) y CS app (inferior)

En la Fig. 74 se puede observar la parte correspondiente del consumidor final en un escenario IDES básico completo, en el cual se aprecia la MS app reproduciendo contenido broadcast en el terminal híbrido (la TV), mientras que hay cuatro dispositivos secundarios ejecutando CS apps, reproduciendo contenido audiovisual que proporciona diferentes puntos de vista del mismo evento, a través de la tecnología DASH (adquirido a través de la red broadband).



Fig. 74 Escenario IDES básico formado por un dispositivo principal (ejecutando la MS app) y 4 dispositivos secundarios (ejecutando la CS app)

Como se puede observar en las anteriores figuras, se ha superpuesto el número de fotograma en el contenido, con el fin de comprobar de una forma rápida y visual el nivel de sincronización adquirido con los mecanismos propuestos. Esta superposición se ha llevado a cabo durante la codificación del contenido, mediante el uso de la herramienta *ffmpeg*. De manera similar, para comprobar visualmente el cambio entre calidades en el contenido adaptativo, además del número de fotograma, en el contenido broadband (DASH) se ha añadido información superpuesta sobre el bitrate y la resolución del contenido.

La Tabla 13 y la Tabla 14, muestran las principales características y especificaciones de los contenidos híbridos generados y de los dispositivos utilizados en la evaluación, respectivamente.

Tabla 13 Contenidos híbridos utilizados en el escenario IDES básico

Tecnología	Contenido	Codificación
Broadcast (MPEG2-TS)	Vídeo	H.264, 1920x1080, 25fps (4 canales, 30Mbps)
	Audio	MPEG-4 AAC 140Kbps (2 canales, 48KHz)
Broadband (DASH)	Perfil de vídeo	H.264, 25 fps, duración de los segmentos: 3s
		Calidad 1 (Q1): 426x240, 700Kbps
		Q2: 640x360, 1000Kbps
		Q3: 854x480, 2000Kbps
		Q4: 1280x720, 4000Kbps
	Q5: 1920x1080, 6000Kbps	
	Audio	MPEG-4 AAC 128Kbps (2 canales, 48KHz)

Tabla 14 Dispositivos utilizados en el escenario IDES básico

Dispositivo	Descripción
PC (broadcaster)	Intel Core i7-6700 @ 3.40GHz, 8GB RAM, SSD 240GB, Windows 10, DVB-T PCI Card: DekTec DTA-2111
PC (servidor multimedia y NTP)	Intel Xeon E5420 @ 2.50GHz x8, 8GB RAM, HDD 200GB, Ubuntu 14.04, Apache Server v2.4.7, Fast Ethernet embedded card.
Switch/Router	TP-Link AC1900 Wireless Dual Band Gigabit Router. IEEE 802.11ac.
Dispositivo principal ejecutando la MS app (Set Top Box, STB)	PC1: Intel Celeron 1037U @ 1.80GHz x2, 4GB RAM, HDD 150GB, Ubuntu 14.04; Fast Ethernet embedded card; WiFi card IEEE 802.11 b/g/n; DVB-T USB card: Hauppauge! Nova-T Stick 3; TV LG 32LF592U (32").
Dispositivo secundario ejecutando la CS app	PC1, Samsung Galaxy Tab S (IEEE 802.11ac 10" Android 7 tablet), Samsung S5 (5" Android 7 smartphone)

Respecto a parámetros configurables del mecanismo de sincronización, por un lado, se ha ajustado el nivel de asincronía máximo tolerable entre la reproducción de la aplicación MS y la de la aplicación CS a ± 80 ms. Este valor es equivalente a una diferencia de 2 fotogramas cuando el contenido reproducido tiene una tasa de 25fps. Aunque este valor exceda los límites para considerar la sincronización como precisa a nivel de trama, encaja dentro de los límites tolerables de sincronización para este tipo de escenarios [163] y proporciona unos resultados de rendimiento de los dispositivos satisfactorios (tal y como muestra en este capítulo) y unos resultados de percepción de sincronización por parte de los usuarios satisfactorios (tal y como se muestra en el Capítulo 7, acerca de la evaluación subjetiva de la plataforma).

Por otro lado, la frecuencia de envío de información temporal, desde la MS app a la CS app, se ha ajustado a un envío cada 15 fotogramas, equivalente a un envío cada 0.6 segundos al tratarse de contenidos con una tasa de 25fps en este escenario.

El fichero RMCF generado para este caso de uso se puede observar en la Fig. 75. Para cada contenido complementario, se ha incluido en cada etiqueta MEDIA la marca de tiempo inicial (parameter *temi_init*) como propiedad para facilitar el proceso de sincronización con el contenido reproducido por la MS app. También se puede observar la inclusión de un servidor NTP en la etiqueta CLOCK, que se utilizará como referencia de tiempo por parte de todos los dispositivos involucrados, con el fin de poder alcanzar un nivel de sincronización híbrida IDEs satisfactorio. En este caso, se ha utilizado el servidor NTP perteneciente a la *Entidad Pública Empresarial Red.es*, ubicado en Valencia (IPv4 147.156.7.50).

```
<?xml version="1.0" encoding="UTF-8"?>
<Related Media Contents File>
  <MEDIA id="1" media_type="AV" media_format="h264/aac"
    metadata="front_camera/spanish"
    temi_init="3699255471291907022">
    <source protocol="http/dash"
      uri="http://IP_Server/multicam_scenari0/cam1/stream.mpd"/>
    <source protocol="rtsp" uri="rtsp://IP_Server:8551/test"/>
  </MEDIA>
  <MEDIA id="2" media_type="AV" media_format="h264/aac"
    metadata="front_left_camera/spanish"
    temi_init="3699255471291907022">
    <source protocol="http/dash"
      uri="http://IP_Server/multicam_scenari0/cam2/stream.mpd"/>
    <source protocol="rtsp" uri="rtsp:// IP_Server:8552/test"/>
  </MEDIA>
  <MEDIA id="3" media_type="AV" media_format="h264/aac"
    metadata="front_right_camera/spanish"
    temi_init="3699255471291907022">
    <source protocol="http/dash"
      uri="http://IP_Server/multicam_scenari0/cam3/stream.mpd"/>
    <source protocol="rtsp" uri="rtsp:// IP_Server:8553/test"/>
  </MEDIA>
  <MEDIA id="4" media_type="AV" media_format="h264/aac"
    metadata="right_side_camera/spanish"
    temi_init="3699255471291907022">
    <source protocol="http/dash"
      uri="http://IP_Server/multicam_scenari0/cam4/stream.mpd"/>
    <source protocol="rtsp" uri="rtsp:// IP_Server:8554/test"/>
  </MEDIA>
  <CLOCK id="1" protocol="ntp" media_type="time"
    media_format="64_bit_ntp_time"
    metadata="" uri="147.156.7.50"/>
  <LASTUPDATE protocol="http" media_type="time"
    metadata="" format="dd/mm/yyyy-hh:mm:ss"
    value="11/04/2017-11:20:00"/>
</Related Media Contents File>
```

Fig. 75 Fichero RMCF utilizado para el Escenario básico IDEs

6.2.2 METODOLOGÍAS DE EVALUACIÓN EMPLEADAS

En escenarios de reproducción de contenidos híbridos en modo mosaico o en modo PiP, aunque se pueda percibir el nivel de sincronización entre los procesos de reproducción de las MS y CS apps; e incluso ciertos “ecos” provocados al escuchar varios audios a la vez, es difícil valorar, de manera precisa, cuál es el nivel de precisión de la sincronización alcanzada. Un modo práctico de comprobar dicho nivel consiste en reproducir en ambas y al mismo tiempo, el mismo contenido recibido tanto por la red broadcast como por la red broadband. Si ambos contenidos se cortan (*cropping*) por la mitad, y las aplicaciones MS y CS reproducen cada una de las mitades cortadas y ambos reproductores se sitúan juntos, se puede valorar si el resultado emula un único contenido (se percibe como sincronizado) o no. La Fig. 76 muestra este efecto al juntar dos mitades del mismo contenido de manera sincronizada (sin discontinuidades en el centro) o no (con discontinuidades).

Si el conjunto de los dos vídeos se perciben como un único vídeo y el audio de ambos se solapa (esto es, no se oyen ecos y son indistinguibles), significa que el nivel de sincronización obtenido es satisfactorio. En caso contrario, un desalineamiento entre los vídeos o la existencia de ecos en los audios significará una falta de sincronización perceptible.

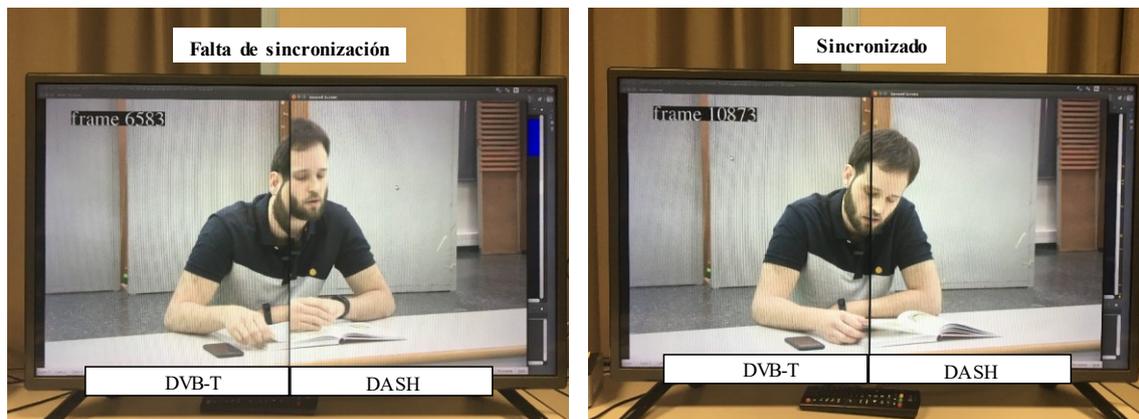


Fig. 76 Mismo contenido broadcast y broadband cortado por la mitad y reproducido por aplicaciones MS (parte izquierda del contenido) y CS (parte derecha del mismo) para comprobar visualmente el nivel de sincronización alcanzado en el terminal híbrido

Una manera más precisa de comprobar visualmente el nivel de sincronización alcanzado consiste en hacer uso de la información de los números de fotograma superpuestos en el contenido. Una vez la plataforma está reproduciendo contenido en múltiples pantallas, se puede grabar un vídeo de la escena o hacer fotos en determinados instantes durante la reproducción. Posteriormente, las fotos pueden ser analizadas o el vídeo grabado puede ser pausado en un momento concreto. Se puede comprobar la diferencia de los números de fotogramas mostrados en las MS y CS apps (estén ejecutándose en el mismo dispositivo que el MS - terminal híbrido- o en diferentes dispositivos). A través de este procedimiento se ha obtenido la Fig. 74. En dicha figura, si se amplía, se pueden observar los siguientes números de fotograma: 5393 en la MS app; 5394 en la CS app de la Tablet izquierda; 5393 en la CS app de la Tablet derecha; y 5394 en las CS apps de ambos smartphones. Este nivel de sincronización implica una asincronía máxima de un fotograma. Observando la Tabla 13, se puede comprobar que el contenido tiene una tasa de 25 fps, por lo que cada fotograma tiene una duración de 40ms. Dicho resultado nos proporciona un análisis inicial acerca del funcionamiento del mecanismo de sincronización híbrida IDEs implementado en el escenario a evaluar.

A continuación, se expone un método más riguroso para evaluar el nivel de sincronización alcanzado entre los procesos de reproducción de las MS y CS apps, de forma mucho más objetiva y automatizada.

En dicho método se ha registrado el nivel de asincronía entre los procesos reproductores de las MS y CS apps. El registro se ha llevado en las aplicaciones CS comparando los *timelines* o líneas temporales (absolutas y relativas) de cada flujo audiovisual, a partir de las marcas de tiempo (*timestamps*) globales insertadas en el descriptor *TEMI timeline* en el contenido DVB-T (que la MS app pasa periódicamente a la CS app) y del valor almacenado en la propiedad *temi-init* en el fichero RMCF para el contenido broadband.

El caso de uso se ha evaluado en diferentes configuraciones (con conexiones cableadas e inalámbricas) y bajo diferentes condiciones de red en el escenario descrito.: en condiciones ideales (sin ningún tipo de limitación de ancho de banda -BW-, retardos o errores de transmisión) y en condiciones más realistas (forzando pérdidas de paquetes, retardos y limitación en el BW). Todos los experimentos han consistido en la realización de 10 sesiones, a partir de las cuales se han obtenido como resultado el valor cuadrático medio (Root Mean Square o RMS) y el intervalo de confianza del 95% (en adelante, I.C. 95%) de los valores de asincronía registrados.

6.2.3 RESULTADOS OBTENIDOS

CONFIGURACIÓN DEL ESCENARIO CON CONDICIONES IDEALES

En este caso, en el escenario no se han configurado ninguna pérdida de paquetes en la transmisión DVB-T, ni tampoco ninguna limitación en el ancho de banda disponible en la red broadband IP utilizada para la adquisición del contenido DASH. Se han llevado a cabo dos experimentos diferentes: a) MS y CS apps en el mismo dispositivo (terminal híbrido); y b) MS y CS apps en dispositivos diferentes, un TV (emulado con un PC), una tablet y un smartphone (múltiples terminales).

MAIN Y COMPANION SCREEN APPS EN EL MISMO DISPOSITIVO (TERMINAL HÍBRIDO)

En este caso, se utilizó únicamente el PC1 (ver Tabla 14), con S.O. Linux Ubuntu 14.04, emulando un terminal híbrido, y se realizaron dos pruebas diferentes, dependiendo del tipo de conectividad a la red broadband del dispositivo: cableada e inalámbrica. Respecto a la conectividad cableada, el dispositivo se conectó al switch del router a través de una conexión Ethernet de 100Mbps. Respecto a la conectividad inalámbrica, el dispositivo se conectó al router (ver Tabla 14) a través de una conexión WiFi de 72Mbps. Por otro lado, el servidor web con el contenido DASH se conectó directamente al switch del router de forma cableada, a través de una conexión Ethernet de 100Mbps. Los resultados obtenidos se presentan en la Tabla 15, en la que puede comprobarse cómo el valor RMS de los valores de asincronía registrados se encuentra por debajo del umbral establecido de 80ms.

Tabla 15 Valor cuadrático medio e I.C. 95% del valor de asincronía obtenidos para el caso ideal en un terminal híbrido (ejecutando las MS y CS apps)

Conexión broadband cableada	Conexión broadband inalámbrica
31,26ms ± 7,94ms	38,51ms ± 10,01ms

MAIN SCREEN Y COMPANION SCREEN EN DIFERENTES DISPOSITIVOS (MÚLTIPLES TERMINALES)

En este experimento, la MS app en el PC1 (ver Tabla 14), emulando un TV, con S.O. Linux Ubuntu 14.04, mientras que la CS app se ejecutó en una tablet de 10" y un smartphone de 5" con S.O. Android versión 7 (ver Tabla 14). Los resultados obtenidos se muestran en la Tabla 16. Tal y como se puede observar, los valores RMS de los valores de sincronía registrados se encuentran por

debajo del umbral establecido de 80ms, y adicionalmente, se observa que la tablet ofrece un mejor rendimiento que el smartphone, lo cual es explicable debido a las especificaciones de ambos dispositivos.

Tabla 16 Valor cuadrático medio e I.C. 95% de los valores de asincronía obtenidos para el caso ideal en entorno multidispositivo (apps MS y CS ejecutándose en dispositivos diferentes)

Tablet	Smartphone
43,61ms ± 11,60ms	56,62ms ± 11,81ms

CONFIGURACIÓN MÁS REALISTA DEL ESCENARIO CON PÉRDIDAS DE PAQUETES, RETARDOS Y LIMITACIÓN EN EL ANCHO DE BANDA

En este caso, se llevaron a cabo tres pruebas diferentes con diferentes configuraciones respecto a pérdidas de paquetes, retardos y limitación en el BW: a) transmisión broadcast con pérdidas y transmisión broadcast ideal; b) transmisión broadcast ideal y transmisión broadband con retardos y limitación en el BW; y c) combinación de las dos anteriores.

Por un lado, tal y como se explicó en el Capítulo 5, en la plataforma se utiliza el software Dektec StreamXpress, encargado de la transmisión broadcast a través de la tarjeta PCI Dektec DTA-2111. Dicho software permite forzar de forma aleatoria la pérdida de paquetes durante la fase de transmisión DVB del contenido. Respecto a la tecnología broadcast DVB-T, cada uno de los paquetes que forman parte del flujo MPEG2-TS transmitido se extiende con un código de protección de errores Reed-Solomon, obteniéndose paquetes de una longitud de 204 bytes. Esta protección en combinación con mecanismos de codificación convolucional y esquemas de modulación apropiados permite garantizar un transporte denominado *quasi-error free* (QEF) de servicios DVB, lo cual significa que, como promedio, sólo ocurre un error no corregible a lo largo de la presentación de una hora de programa (equivalente a un BER de 1×10^{-11}). Teniendo en cuenta que la duración de los vídeos utilizados durante la evaluación es de aproximadamente 2 minutos, se ha configurado una probabilidad de pérdida de paquetes 0.05% (significativamente mayor que el valor medio en QEF) en la transmisión broadcast.

Por otro lado, se ha utilizado el emulador de condiciones de red *netem*¹⁷ para forzar diferentes condiciones de disponibilidad de ancho de banda en la conexión broadband durante los experimentos, además de un retardo de red de 60ms ± 20ms siguiendo una distribución normal (lo cual corresponde con lo que se puede observar en conexiones fijas de larga distancia o redes móviles, por lo que es un valor representativo para una amplia variedad de aplicaciones). Con el objetivo de forzar la adquisición y reproducción de diferentes calidades (es decir, de forzar el *quality switching*) en el reproductor de DASH de la CS app, se ha diseñado un patrón de limitación de BW específico para que, de esta forma, el reproductor se vea obligado a reproducir todas las calidades disponibles del contenido. Concretamente, el BW disponible en la conexión Ethernet del servidor web convencional se ha limitado, en etapas de 25 segundos, a 900kbps, 1700kbps, 2300kbps, 4100kbps y 'sin limitación' (ver los bitrates correspondientes a cada calidad generada para el contenido DASH en la Tabla 13). Adicionalmente, en dicha conexión también se forzó un retardo de red de 60ms ± 20ms (siguiendo una distribución normal).

A continuación, se muestran los resultados para las tres pruebas realizadas, las cuales se han realizado estando tanto las MS y CS apps en el mismo dispositivo (terminal híbrido) como en múltiples dispositivos independientes (múltiples terminales).

¹⁷ <https://wiki.linuxfoundation.org/networking/netem>

MAIN SCREEN Y COMPANION SCREEN APPS EN EL MISMO DISPOSITIVO (TERMINAL HÍBRIDO)

Los resultados obtenidos cuando las MS y CS apps se han ejecutado en el mismo dispositivo (terminal híbrido) se presentan en la Tabla 17, en la que, adicionalmente, se diferencia cuando dicho dispositivo se ha conectado a la red broadband de forma cableada o inalámbrica. Se puede observar que el valor RMS de los valores de asincronía registrados es inferior al umbral configurado de 80ms.

Tabla 17 Valor cuadrático medio e I.C. 95% de los valores de las asincronías registradas en un escenario realista utilizando un único terminal híbrido (ejecutando las MS y CS apps)

Condiciones	Conexión cableada	Conexión inalámbrica
Probabilidad de pérdida de paquete broadcast 0.05%	54,54ms ± 11,63ms	58,94ms ± 13,48ms
Limitaciones de BW y retardos añadidos en la conexión broadband	48,07ms ± 10,30ms	49,34ms ± 11,87ms
Combinación de las dos anteriores	56,59ms ± 12,87ms	60,56ms ± 13,61ms

MAIN SCREEN Y COMPANION SCREEN EN DIFERENTES DISPOSITIVOS (MÚLTIPLES DISPOSITIVOS)

Los resultados obtenidos en este caso se resumen en la Tabla 18. Al igual que en el caso anterior, también se puede comprobar que el valor RMS de los valores de sincronía registrados se mantiene por debajo del umbral de 80ms, aunque, como era de esperar, al tratarse de diferentes dispositivos, son valores más altos que cuando las MS y CS apps se han ejecutado en el mismo dispositivo. De manera similar al caso ideal, cuando la CS app se ejecuta en otro dispositivo distinto al que ejecuta la MS app, se puede observar cómo el smartphone ofrece un peor rendimiento que la Tablet. Tal como se ha indicado anteriormente, ello es debido a que las especificaciones y recursos del smartphone son peores. De hecho, el valor medio obtenido para el smartphone en la prueba en la que se fuerzan tanto pérdidas de paquetes en la transmisión broadcast como una limitación en el BW y retardos en la red broadband, llega a exceder el umbral, si se tiene en cuenta el I.C. 95% obtenido. Sin embargo, los resultados se pueden seguir considerando satisfactorios, teniendo en cuenta las malas condiciones de red forzadas (tanto en la red broadcast como en la red broadband) así como el limitado rendimiento del dispositivo en cuestión.

Tabla 18 Valor cuadrático medio e I.C. 95% del valor de las asincronías registradas en un escenario realista multidispositivo (MS y CS apps ejecutándose en dispositivos diferentes)

Condiciones	Tablet	Smartphone
Probabilidad de pérdida de paquete broadcast 0.05%	61,05ms ± 13,17ms	64,61ms ± 13,14ms
Limitaciones de BW y retardos añadidos en la conexión broadband	51,93ms ± 12,58ms	58,91ms ± 12,20ms
Combinación de las dos anteriores	61,69ms ± 14,10ms	77,88ms ± 15,78ms

Para este caso, la Fig. 77 muestra el registro de los valores de asincronía registrados por la CS app en el smartphone durante una de las sesiones del caso multidispositivo con cierta probabilidad de pérdida de paquetes en la transmisión broadcast.

Concretamente, la gráfica superior izquierda muestra el valor de asincronía registrado durante aproximadamente 180 segundos. Se puede observar cómo inicialmente aparece una etapa transitoria hasta que se alcanzan valores del nivel de sincronización por debajo del umbral configurado. Tal y como se observa en la gráfica inferior izquierda, los niveles de sincronización registrados se mantienen estables por debajo del umbral configurado (± 80 ms), aunque se puede

apreciar que el proceso de reproducción presenta fluctuaciones que pueden provocar que, puntualmente, se registren valores que superan dicho umbral. Durante dicha sesión, se realizó un cambio de contenido (mediante el menú con botones, mostrado en la Fig. 77) aproximadamente en el segundo 134, observable en las gráficas superiores (en la de la derecha se representa una ampliación para que se pueda apreciar mejor). Se puede comprobar cómo, tras una etapa de ajustes (se realizan tres ajustes, marcados con puntos rojos en la gráfica superior derecha), el proceso de reproducción retorna a niveles de sincronización con valores de asincronía por debajo del umbral configurado, lo cual permita que se llegue rápidamente de nuevo a una fase estable en la reproducción del contenido. Cabe resaltar que, el resto de los picos que se aprecian en dicha gráfica son los causantes de los ajustes que se llevan a cabo en el proceso de reproducción. Debido al rendimiento heterogéneo de los dispositivos CS, se decidió implementar en el algoritmo de ajuste del nivel de sincronización, una monitorización de la fluctuación del proceso de reproducción consistente en esperar a que dicho proceso se mantenga estable tras un ajuste previo. De esta manera, se evita llevar a cabo falsas detecciones de asincronía debidas a una fluctuación (irregular, arbitraria e inevitable) en el proceso de reproducción inmediatamente tras realizar un ajuste. Adicionalmente, en la gráfica inferior derecha de la Fig. 77, se muestra el porcentaje en tanto por uno de la frecuencia de registros de los valores de asincronía obtenidos, agrupados en rangos de 20ms.

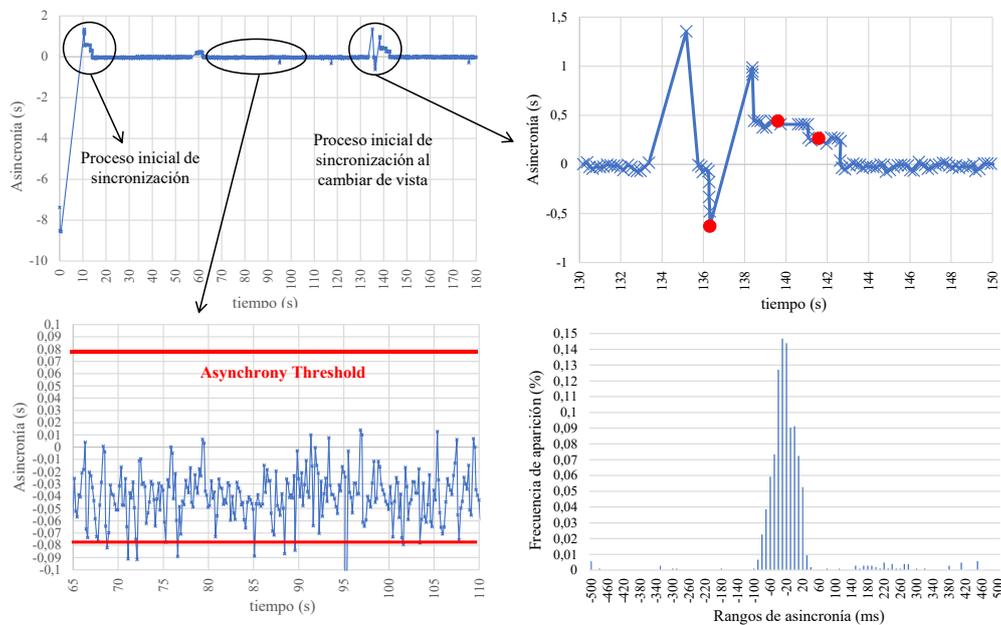


Fig. 77 Valores de asincronía registrados por la CS app durante una sesión de la evaluación objetiva y la distribución de los valores registrados

6.2.4 DISCUSIÓN

Durante todas las pruebas realizadas, el valor RMS de los valores de asincronía registrados se ha mantenido por debajo del umbral establecido (a 80ms absolutos) en todo momento. Al ejecutar la CS app, tras un proceso de sincronización inicial, el nivel de sincronización se ha mantenido dentro de los límites establecidos durante la mayor parte del tiempo. Al seleccionar otro contenido complementario (otro punto de vista de la escena de lectura) en la CS app, el mecanismo de sincronización diseñado permitió alcanzar rápidamente niveles de sincronización satisfactorios mediante ajustes suaves (AMP, +/-15% de la tasa nominal del proceso de reproducción) en el proceso de reproducción. A pesar de existir una cierta fluctuación en los valores de sincronización registrados, dichos valores se han mantenido por debajo del umbral salvo en situaciones ocasionales, principalmente debido a las pérdidas de paquete, a la limitación

eventual del BW, al cambio de contenido o al propio inicio del proceso de reproducción de la CS app. Sin embargo, dichas fluctuaciones no han supuesto un impacto negativo en la percepción por parte de los usuarios de la fluidez del proceso de reproducción, tal y como se expondrá en el siguiente capítulo, correspondiente a la evaluación subjetiva.

Por un lado, tal y como puede observarse en la Tabla 15 y en la Tabla 17, los valores RMS de asincronía registrados han sido más altos cuando el terminal híbrido se ha conectado de manera inalámbrica en la red de área local, al compararlo con el caso en el que el terminal híbrido se conecta a la misma a través de un cable ethernet.

Por otro lado, los valores RMS de asincronía registrados también han sido superiores cuando la CS app se ha ejecutado en un dispositivo secundario, comparado con el caso en el que se ejecuta en el terminal híbrido. Adicionalmente, también se aprecia que cuando se utiliza un smartphone como CS app, el valor RMS es superior a cuando se emplea una Tablet.

Finalmente, los peores resultados (mayores valores de asincronía), como era de esperar, se corresponden con la configuración realista al utilizar múltiples dispositivos independientes ejecutando la CS app. Adicionalmente, se puede observar cómo, en dicha configuración, los valores de asincronía obtenidos cuando se fuerzan pérdidas de paquetes en la transmisión broadcast son similares a cuando se fuerza una combinación de pérdidas de paquetes en la red broadcast y limitaciones de BW en la red broadband. Sin embargo, cuando únicamente hay limitación de BW y retardos en la red broadband, el nivel de sincronización alcanzado no se ve tan afectado. Esto significa que los errores o pérdidas de paquetes en la red broadcast afectan de una manera más significativa al rendimiento de los mecanismos de sincronización diseñados que cuando solo existen limitaciones o problemas relacionados con la red broadband. En este último caso, dichas limitaciones se gestionan de forma satisfactoria en la conexión HTTP/TCP y mediante técnicas adaptativas adecuadas de buffering y de selección de calidades del contenido, implementadas en el reproductor DASH de la CS app.

Por lo tanto, tal y como se muestra en esta sección, la solución de sincronización híbrida diseñada en esta Tesis Doctoral entre los procesos de reproducción de las MS y CS apps, permite que los valores medios de asincronía permanezcan por debajo del umbral configurado, tanto en escenarios mono- como multidispositivo, cableados o inalámbricos, incluyendo y forzando determinadas condiciones de red (por ejemplo, pérdida de paquetes broadcast, o latencia y retardos de red broadband), e incluyendo dispositivos con recursos de procesamiento limitados (por ejemplo, smartphones).

6.3 ESCENARIO INTER-DISPOSITIVO INMERSIVO

En esta sección se describe la evaluación objetiva realizada en el escenario IDES con contenido de vídeo omnidireccional complementario, accesible a través de la red broadband. En este caso, la aplicación CS ejecutada en los dispositivos secundarios, que recrea un escenario inmersivo de consumo de contenidos, consistente en un salón virtual con un menú con varias opciones de selección de contenidos y un reproductor 2D y omnidireccional (ver Fig. 78), explicado en el Capítulo 5.

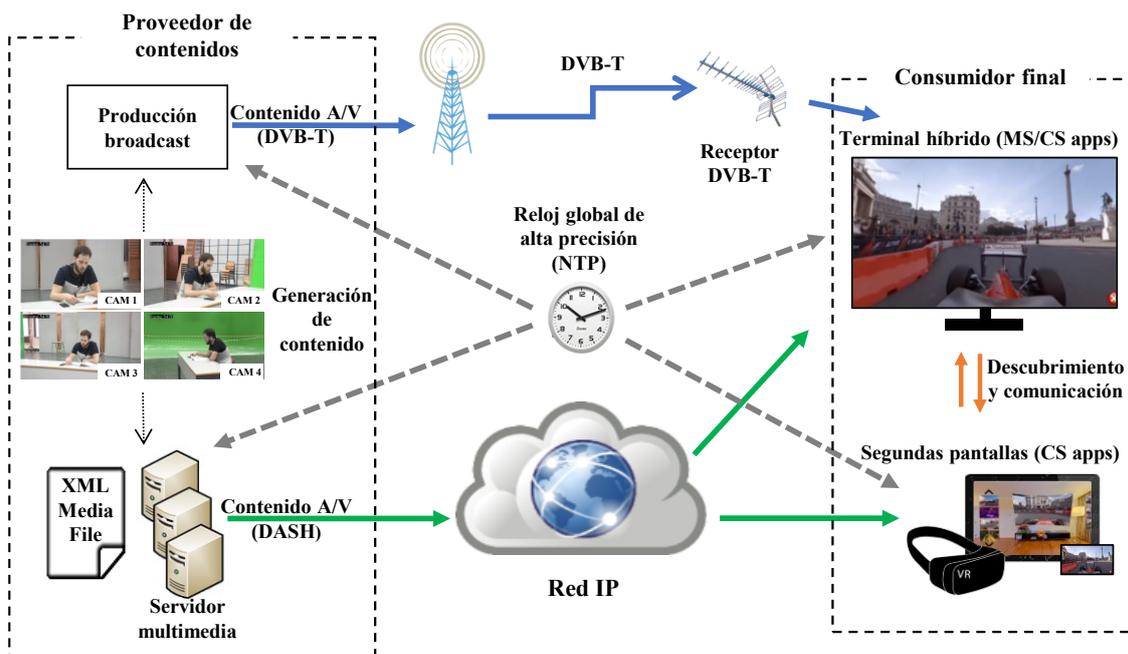


Fig. 78 Escenario ides implementado en la plataforma para la evaluación objetiva de la solución ides inmersiva (incluyendo HMD)

6.3.1 ESCENARIO DE EVALUACIÓN

Los dispositivos involucrados en este escenario están listados en la Tabla 19. Además, las principales características del contenido generado se muestran en la Tabla 20. Se puede observar que los dispositivos secundarios utilizados para esta evaluación cuentan con mejores prestaciones que los utilizados en la evaluación del escenario IDES básico. Esto se debe a que se necesitan dispositivos con mejores especificaciones para poder reproducir contenido de vídeo omnidireccional.

Tabla 19 Dispositivos involucrados en el escenario IDES inmersivo

Dispositivo	Rol
Intel Core i7-6700 @ 3.40GHz CPU, 8GB RAM, W10 with DTA-2111 modulator	Broadcaster
Intel Xeon E5420, 8GB RAM, Ubuntu 14.04, Fast Ethernet 100	Broadband Media Server
PC: Celeron N3050 @ 1.60GHz CPU, 4GB RAM, Ubuntu 14.04 with DVB USB rtl2832, connected to an LG 32LF592U 32" Smart TV (MS) via HDMI	Hybrid Terminal + MS
Tablet Samsung Galaxy Tab S4	CS #1
Smartphone Samsung Galaxy S5	CS #2
Smartphone Samsung Galaxy S6	CS #3
Smartphone Samsung Note 9	CS #4
Portátil w/ Intel Core i7-7700HQ @ 2.8GHz, 16GB RAM, Windows 10 OS	CS #5
TP-Link AC1900	Router inalámbrico

Tabla 20 Características del contenido omnidireccional

Característica	Valor configurado
Codificación del contenido	H.264 (sin tiles) + AAC
Tasa de fotogramas	25fps
Resolución del contenido broadcast (MPEG2-TS)	1920x1080px
Resolución del contenido broadband (DASH), con segmentos de vídeo de 3"	7680x3840px 3840x1920px 1920x960px 960x480px 480x240px

Para el contenido omnidireccional se ha escogido la codificación h.264 y no otro más eficiente, como por ejemplo, HEVC, porque en el momento de la evaluación este tipo de codificaciones todavía no estaban soportadas por la mayoría de navegadores (solamente por aproximadamente un 17% de los navegadores [164]). En ese momento, H.264 era la codificación con mayor soporte (aproximadamente por un 98% de los navegadores [165]).

La parte izquierda de la Fig. 79 muestra el aspecto del salón virtual, donde se reproduce el contenido 2D, implementado en el dispositivo secundario (CS app). A la parte derecha, se muestra el aspecto del reproductor de contenido omnidireccional en dicho dispositivo.



Fig. 79 Aspecto del salón virtual visto en pantalla plana (izquierda superior) y en HMD (izquierda inferior). Aspecto del reproductor omnidireccional visto en pantalla plana (derecha superior) y en HMD (derecha inferior).

Para indicar la disponibilidad de este tipo de contenido en la plataforma, en el fichero RMCF se ha añadido la información necesaria utilizando las etiquetas explicadas en el Capítulo 4, tal y como se muestra en la Fig. 80.

```
<HybridMediaContentsFile>
  <MEDIA id="1" media_type="360AV" media_format="DASH-SRD"
  metadata="360_camera/english" temi_init="3760076589162762510">
    <source projection="ERP" protocol="http/dash" tiled="false"
    uri="http://W.X.Y.Z/360/ERP/stream360.mpd"/>
    <source projection="CMP" protocol="http/dash" tiled="false"
    uri="http://W.X.Y.Z/360/tiled/CMP/stream360tiled.mpd"/>
  </MEDIA>
  <MEDIA id="2" media_type="AV" media_format="h264/aac"
  metadata="2D_front_camera/english" temi_init="3760076589162762510">
    <source protocol="http/dash" uri="http://W.X.Y.Z/2D/stream2D.mpd"/>
    <source protocol="rtsp" uri="rtsp://W.X.Y.Z:1234/2d"/>
  </MEDIA>
  [...]
</HybridMediaContentsFile>
```

Fig. 80 Etiquetas y parámetros específicos para señalar la disponibilidad del contenido omnidireccional en el fichero RMCF

Se ha empleado un contenido denominado *F1&Roller Coaster*, cuyas características están listadas en la Tabla 21, y que cuentan con todo tipo de escenas (lentas y rápidas). En particular, este contenido se ha generado a partir de dos contenidos omnidireccionales independientes, cuya localización está reflejada en la Tabla 21. Se ha optado por la opción de concatenar y generar un contenido personalizado debido a la falta de contenidos omnidireccionales disponibles con una

duración y una temática atractiva para el objetivo que se expone en los Capítulos 6 y 7 de esta Tesis Doctoral (relacionadas con la evaluación objetiva y subjetiva de la plataforma, respectivamente).

El contenido generado consta de dos partes, en la primera de ellas se muestra un circuito urbano de F1, con una cámara omnidireccional que ha sido embarcada en el mismo vehículo. En dicho contenido hay tanto escenas rápidas en las que el vehículo acelera como lentas, en las que el vehículo realiza una serie de desaceleraciones y giros. En la segunda parte, se presenta un viaje en montaña rusa, con una cámara omnidireccional embarcada en uno de los vagones de la atracción. En esta segunda parte, también se presentan escenas rápidas y lentas, como en las caídas en picado o subidas de rampas, respectivamente.

Una vez concatenados ambos contenidos, se han generado, adicionalmente, 3 vistas tradicionales, obtenidas del recorte de las regiones del contenido omnidireccional correspondientes a: 1) vista frontal; 2) vista lateral izquierda; y 3) vista lateral derecha. Para estas 3 vistas adicionales, se han generado diferentes calidades siguiendo las características expuestas en la Tabla 21.

Tabla 21 Características del contenido generado para la evaluación

Título	Duración (s)	Codificación	Vistas 2D	Vistas 360	Proyección 360	Fuentes
F1 & Roller Coaster	360	H.264 + AAC	3	1	CMP	https://www.youtube.com/watch?v=fQoVFraB0nc&t=9s https://www.youtube.com/watch?v=8lsB-P8nGSM&t=89s

6.3.2 METODOLOGÍA DE EVALUACIÓN EMPLEADA

Se ha evaluado el funcionamiento de las soluciones de sincronización en este escenario, se han utilizado diferentes dispositivos secundarios que ejecutarán la CS app (ver Tabla 19). En este caso, al igual que en la evaluación del escenario IDES básico, se ha establecido un umbral máximo de asincronía tolerable entre los procesos de reproducción de la MS app y la CS app de ± 80 ms. La probabilidad de pérdida de paquetes en la transmisión broadcast se ha configurado al 0.1%, y el retardo de la red broadband se ha forzado para que esté comprendida entre $60\text{ms} \pm 20\text{ms}$ siguiendo una distribución normal. En este caso, solamente se ha llevado a cabo la evaluación objetiva bajo estas condiciones, omitiendo la evaluación en un escenario ideal.

Para validar el mecanismo de sincronización híbrida IDES cuando se está reproduciendo contenido en el dispositivo principal (ejecutando la MS app) y contenido omnidireccional en los dispositivos secundarios (ejecutando la CS app), se ha llevado a cabo un registro, en estos últimos, de los valores de asincronía obtenidos durante 10 sesiones de aproximadamente 3 minutos, tiempo suficiente para que el dispositivo secundario alcanzara un estado de sincronización estable.

6.3.3 RESULTADOS OBTENIDOS

A continuación, la Tabla 22 resume los valores cuadráticos medios (RMS) de los valores de asincronía registrados y el intervalo de confianza del 95%.

Tabla 22 Valor cuadrático medio (RMS) e I.C. 95% de los valores de asincronía registradas

Dispositivo secundario ejecutando la CS app	RMS \pm I.C. 95% (ms)
Portátil	59 \pm 7
Samsung Note 9	60 \pm 2
Samsung Galaxy Tab S4	62 \pm 5

Samsung Galaxy S6	66 ± 3
Samsung Galaxy S5	67 ± 3

Tal y como se puede observar, el valor RMS de los valores de asincronía registrados se encuentra por debajo del umbral configurado (80ms). A pesar de no haber diferencias significativas, el mejor resultado se ha obtenido para el portátil. Por otro lado, se puede observar cómo, en general, los resultados son mejores cuando el dispositivo utilizado es más moderno y cuenta con mejores prestaciones en cuanto a mejor rendimiento, como es el caso del portátil, el smartphone Samsung Note 9 o la Tablet Samsung Galaxy Tab S4. Sin embargo, en los smartphones Samsung Galaxy S5 y S6 se obtuvieron peores resultados (comparándolos con los anteriores dispositivos secundarios). Las Fig. 81 y Fig. 82 muestran la frecuencia de repetición de los valores de asincronía registrados y la distribución de frecuencia acumulada para cada uno de los dispositivos secundarios, respectivamente.

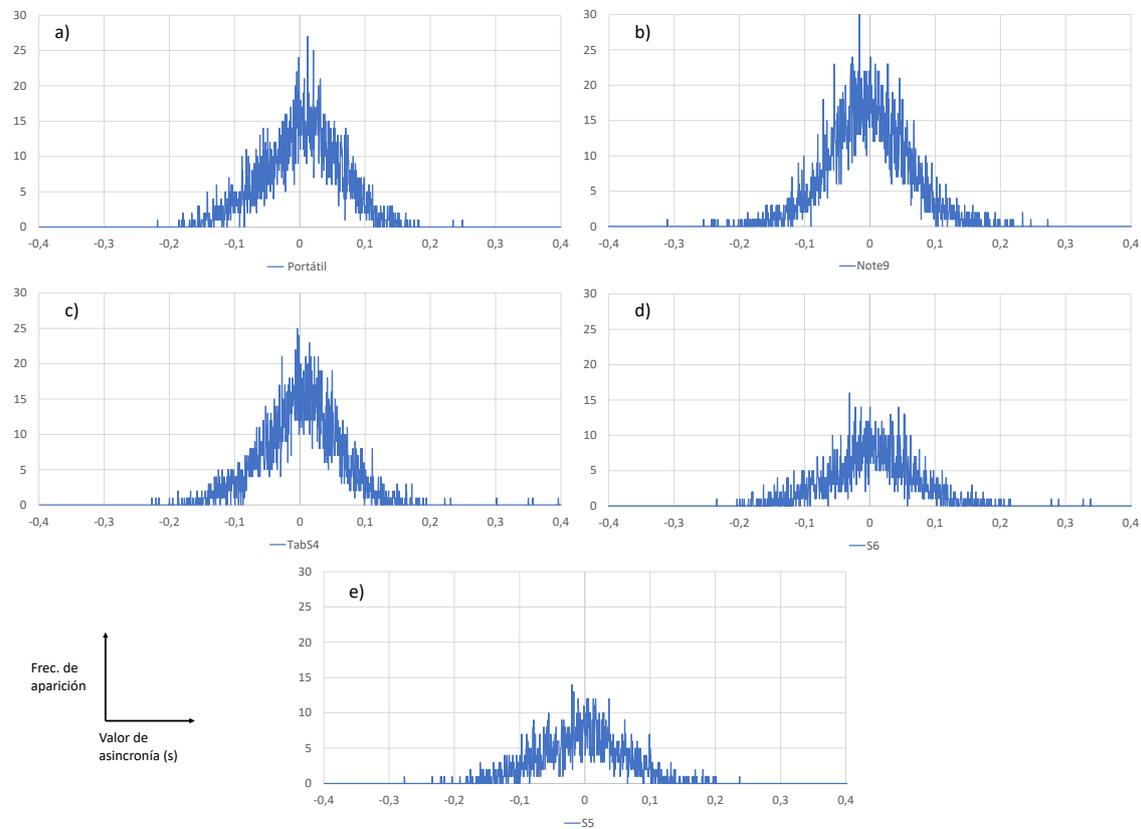


Fig. 81 Valores de asincronía registrados por cada dispositivo secundario y número de coincidencias durante la sesión

En la Fig. 81 se puede observar cómo los dispositivos que mejores niveles de sincronización alcanzan son el portátil, el Samsung Note 9 y la Galaxy Tab S4, al tener un amplio número de repeticiones en valores de asincronía cercanos al 0. Sin embargo, se puede ver cómo esta "altura" decae para los dispositivos smartphones Samsung Galaxy S5 y S6 (aún más en el primero).

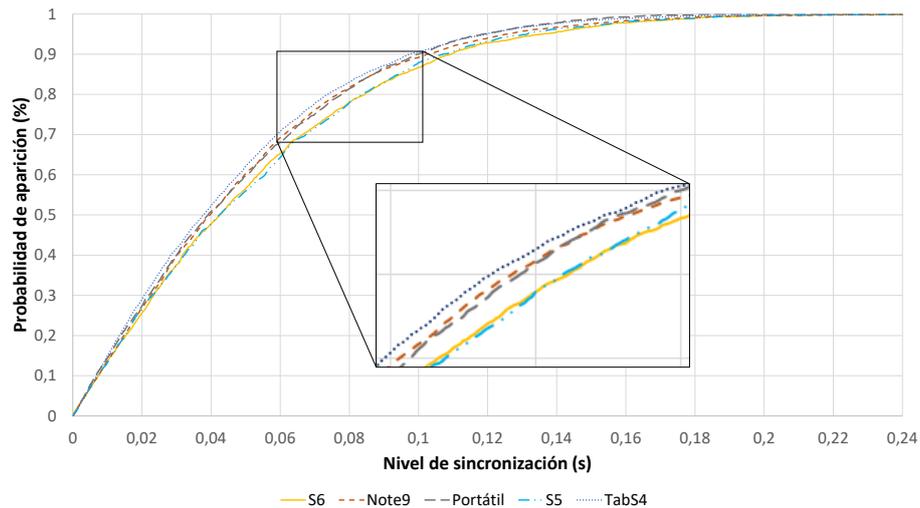


Fig. 82 Función de frecuencia acumulada de los valores absolutos de las asincronías registradas en cada dispositivo secundario

La Fig. 82 complementa los resultados mostrados en la Fig. 81, puesto que en ella se puede ver cómo la pendiente es más pronunciada en los tres dispositivos que ofrecen mejor rendimiento en cuanto al nivel de sincronización alcanzada. En este tipo de figuras, los dispositivos que presentan un gráfico con una mayor pendiente implican un mejor rendimiento en términos del nivel de sincronización híbrida IDES alcanzado.

Según los valores de asincronía obtenidos, los dispositivos S5 y S6 tienen aproximadamente un 78% de probabilidades de registrar un valor de asincronía con un valor absoluto menor o igual a 80ms, mientras que los otros dispositivos cuentan con una probabilidad de aproximadamente el 83% para registrar este mismo valor de nivel de sincronización.

A continuación, se muestran los resultados obtenidos en una sesión adicional, de aproximadamente 10 minutos de duración, en la que se han registrado los valores de asincronía de todos los dispositivos secundarios. En la Fig. 83 se muestran dichos valores para cada uno de ellos. Se puede observar en la misma cómo todos los dispositivos secundarios involucrados registran valores de asincronía que se encuentran dentro de los límites máximos establecidos por el umbral de ± 80 ms. En la Fig. 84, se muestra en detalle el rendimiento ofrecido por el Smartphone Samsung Galaxy S5, con el objetivo de analizar el comportamiento del dispositivo secundario con los peores resultados. Se puede comprobar cómo, cuando los valores de las asincronías registradas en este dispositivo superan el umbral del nivel de sincronización máximo tolerable, se llevan a cabo ajustes que vuelven a situar el valor de las asincronías dentro de los límites configurados. Sin embargo, a pesar de contar con resultados promedio satisfactorios, se puede deducir que la reproducción sufre fluctuaciones que pueden afectar a la QoE experimentada por los usuarios.

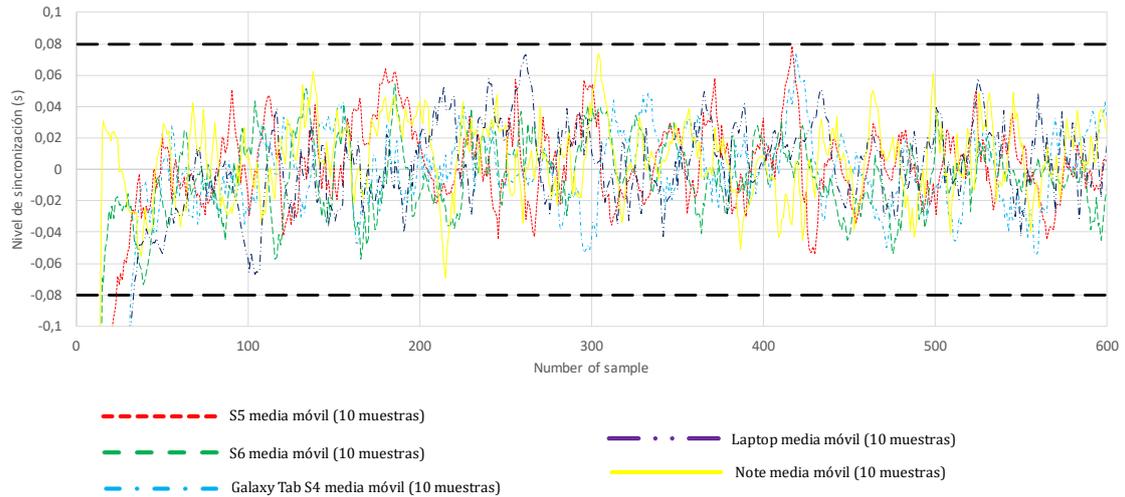


Fig. 83 Valores de asincronía registrados por los dispositivos secundarios durante una sesión de 10 minutos

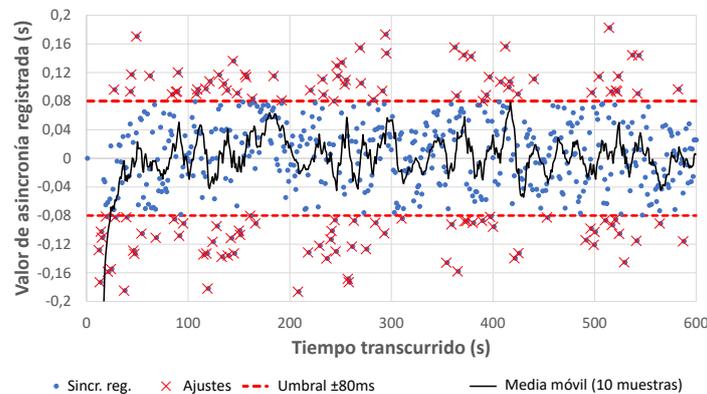


Fig. 84 Valores de asincronía y ajustes registrados en el dispositivo secundario smartphone Samsung Galaxy S5 durante una sesión adicional de 10 minutos (se incluye una media móvil tomando grupos de 10 muestras)

6.3.4 DISCUSIÓN

En vista de los resultados obtenidos en la evaluación objetiva de este escenario, se puede concluir que los mecanismos de sincronización híbrida IDES implementados en los dispositivos principal (ejecutando MS app) y secundarios (ejecutando la CS app) funcionan satisfactoriamente, puesto que los valores medios de asincronía obtenidos se encuentran por debajo del umbral configurado de ± 80 ms. Sin embargo, sí que se observa que el rendimiento general del dispositivo secundario depende directamente de sus características y especificaciones técnicas. Se ha podido comprobar cómo con dispositivos menos recientes (como por ejemplo, el smartphone Samsung Galaxy S5), el comportamiento del mecanismo de sincronización híbrida IDES implementado se ve afectado y se obtienen peores resultados, aunque, en todo caso, siguen siendo valores por debajo del umbral configurado.

Por tanto, se puede concluir que los resultados obtenidos en esta evaluación también han sido satisfactorios.

6.4 ESCENARIO INTER-DESTINATARIO

En esta sección se va a evaluar los mecanismos de sincronización híbrida IDES e IDMS presentados en esta Tesis Doctoral para un caso de uso multidestinatario y multidispositivo, como es el caso de uso de TV Social TV multidispositivo, con contenido complementario multivista (o multicámara) y un servicio de chat privado como medio de interacción social. En este caso, la MS app, además de incluir los mecanismos de la solución de sincronización híbrida IDES propuesta, también incluirá los mecanismos de la solución de sincronización IDMS propuesta. Esta última sólo afecta a la sincronización de la reproducción del contenido principal en los dispositivos que ejecutan la MS app en cada destino (normalmente en el TV o PC que lo emula). Sin embargo, los ajustes de dicha reproducción afectarán indirectamente a su sincronización con la reproducción de contenido complementario en las CS apps de los dispositivos secundarios involucrados en el escenario. Es por ello, que se necesita una solución de sincronización híbrida IDES que se encargará de garantizar la sincronización entre los procesos de reproducción de los contenidos principal y secundarios. La Fig. 85 muestra el aspecto general del escenario a evaluar. Las MS y CS apps utilizadas en el escenario incluye la implementación de la solución de sincronización híbrida IDES propuesta en la presente Tesis Doctoral. Ambos tipos de sincronización (IDES e IDMS) son evaluadas en esta sección.

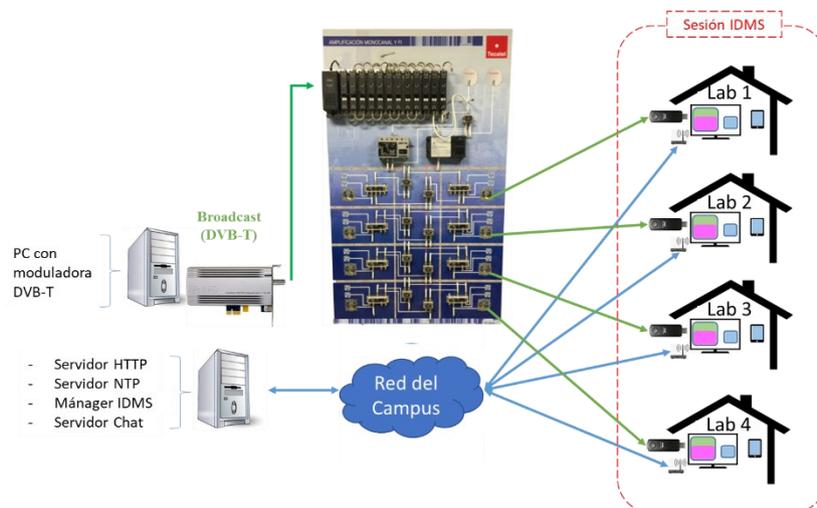


Fig. 85 Visión general del escenario inter-destinatario

6.4.1 ESCENARIO DE EVALUACIÓN

El escenario implementado involucra a los dispositivos listados en la Tabla 23, siendo el contenido de vídeo el mismo que se empleó en el escenario básico IDES y, además, incluye flujos de texto (codificado en ASCII) provenientes de la aplicación de chat integrada en las CS apps (descrita en el Capítulo 5) para habilitar la comunicación mediante mensajes de texto entre usuarios remotos de una misma sesión compartida.

Tabla 23 Dispositivos involucrados en la evaluación del escenario IDMS

Dispositivo	Características
PC _B (broadcaster)	Intel Core i7-6700 @ 3.40GHz, 8GB RAM, SSD 240GB, Windows 10
Transmisor DVB-T	DVB-T PCI Card: DekTec DTA-2111 + DekTec StreamXpress SW
Distribución DVB	Tecatel SMATV Training Panel (con amplificadores, divisores de señal, etc.).
PC _S (servidor multimedia, NTP y Manager IDMS)	Intel Xeon E5420 @ 2.50GHz x8, 8GB RAM, Ubuntu 14.04, Apache Server v2.4.7, Fast Ethernet embedded card.
Routers	4xR: TP-Link AC1900 Wireless Dual Band Gigabit. IEEE 802.11ac.

Hybrid Terminal (Set Top Box, STB)	4xPC _H : Intel Celeron 1037U @ 1.80GHz x2, 4GB RAM, Ubuntu 14.04; Fast Ethernet embedded card; WiFi card IEEE 802.11 b/g/n
Receptor DVB	DVB-T USB card.with RTL2832u chipset
TVs x4 (monitores)	TV1: LG 32LF592U (32"); TV2: Panasonic TH42PX80E (42"); TV3 & TV4: Samsung Smart TV UE48H6200AWXXC (48").
Dispositivos secundarios (apps CS)	2x Samsung Galaxy Tab S (IEEE 802.11ac 10" tablet) 1x Samsung S5 (IEEE 802.11ac 5" smartphone) 1x Samsung S6 (IEEE 802.11ac 5" smartphone)

A diferencia de los dos escenarios anteriores, en este escenario se emulan cuatro "hogares", que están situados en cuatro laboratorios diferentes del Campus de Gandía de la UPV (cada uno con su propia red de área local privada), emulando así múltiples destinatarios remotos. Los dispositivos principales que ejecutan la MS app son cuatro mini-PC que comparten las mismas especificaciones, incluyendo una tarjeta capturadora de TV (DVB-T) y conectados cada uno a un TV. Los dispositivos secundarios empleados en este escenario son las dos tablets Samsung Galaxy Tab S y dos smartphones Samsung Galaxy modelos S5 y S6. La Fig. 86 muestra un ejemplo de un escenario IDMS con dos destinatarios u hogares. Para poder capturar la escena, cada MS con sus respectivos CS se han ubicado sobre la misma mesa de laboratorio. En cada uno de los destinatarios, la MS app reproduce el contenido principal en dos dispositivos principales (TVs) y los dispositivos secundarios de cada destinatario pueden reproducir bien contenido complementario o bien ejecutar una herramienta de chat como medio de interacción entre usuarios remotos

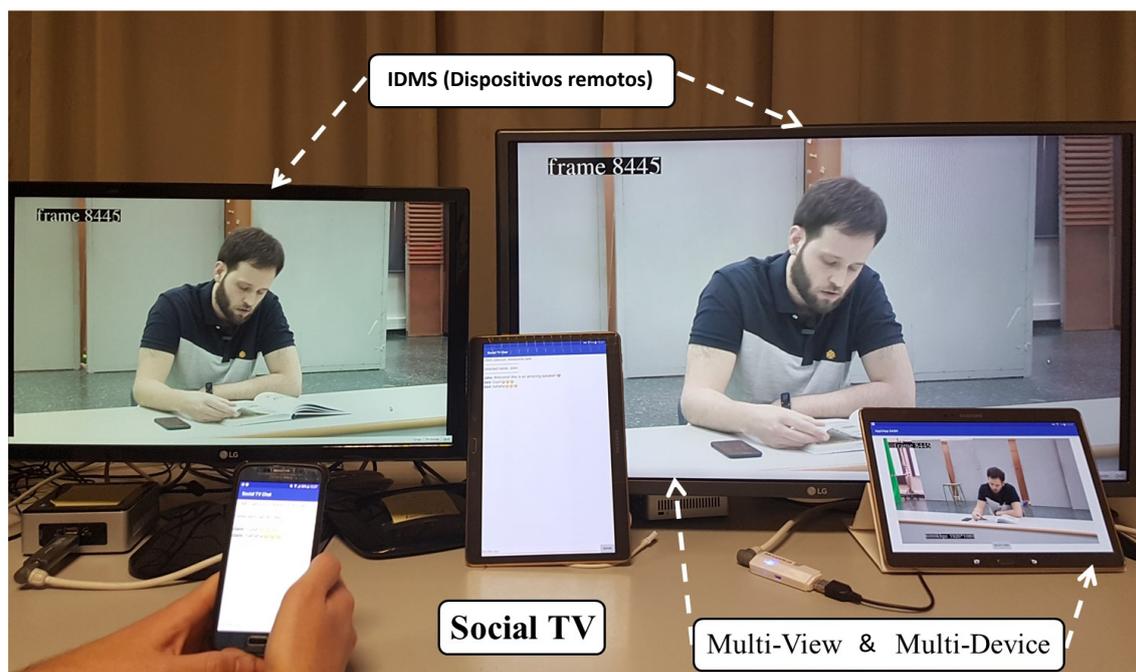


Fig. 86 Ejemplo de escenario IDMS con dos destinatarios independientes.

Respecto a la transmisión de contenido y gestión de información a través de la red broadband, el servidor multimedia, el servidor de chat (necesario si el esquema de control adoptado es SMS al tratarse de un esquema de control centralizado, y opcional si el esquema de control adoptado es DCS), el servidor NTP y el módulo Mánager de IDMS (si el esquema de control adoptado es SMS), se ubican en la misma máquina definida en la Tabla 23 como PC_S (servidor multimedia, NTP y Manager IDMS).

Respecto a la transmisión del contenido a través de la red broadcast, por un lado, tal y como se puede observar en la Fig. 85, se ha utilizado un entrenador de SMATV, emulando un edificio de 4 plantas con 8 viviendas. La salida de la tarjeta moduladora del PC broadcaster con la señal broadcast DVB-T se inyecta a la entrada de la cabecera de amplificación de señal y, posteriormente, se distribuye de manera individual a cada uno de los hogares. Por otro lado, para emular el retardo en la recepción del contenido broadcast en cada una de las apps MS ejecutándose en los dispositivos principales de los cuatro hogares, se ha forzado una latencia fija de recepción diferente en cada uno de ellos: retardo en hogar 1 = 0s; retardo en hogar 2 = 1.5s; retardo en hogar 3 = 3s; y retardo en hogar 4 = 6s. Estos valores se han seleccionado a partir de los resultados presentados en [19], donde se expone que el retardo máximo observado durante una transmisión broadcast es de 6s.

Respecto a los parámetros configurables con el mecanismo de sincronización IDMS, se ha configurado la frecuencia de envío de mensajes *IDMS Report* cada 2s. Además, el umbral de sincronización *session async* (valor de asincronía máximo tolerable en una sesión IDMS, definido en el Capítulo 4) se ha configurado con un valor de ± 160 ms, con el objetivo de no exceder en ningún momento los ± 400 ms, que es el umbral recomendado en escenarios relacionados con TV Social [95]. Tras diferentes pruebas con diferentes valores (menores de ± 400 ms), ± 160 ms fue el mínimo valor con el que la plataforma no se veía afectada en su funcionamiento. Con este valor, si los procesos de reproducción de las MS apps, partícipes de una sesión de consumo de contenidos IDMS se encuentran con un nivel de asincronía por debajo del umbral *session async*, en el peor de los casos consumirán el contenido con 4 fotogramas de diferencia (según el contenido a 25fps generado para la evaluación). Respecto al mecanismo de sincronización híbrida IDMS, el umbral entre dispositivos también se ha configurado en ± 80 ms.

Para este escenario IDMS, el fichero RMCF empleado se muestra en la Fig. 87. A diferencia del fichero RMCF del escenario IDMS básico, en este se incluye la etiqueta <IDMS>, donde se indica la ubicación del IDMS Manager para el caso en el que el esquema de control sea centralizado (SMS).

```
<Related Media Contents File>

<MEDIA id = "media1" media_type = "AV" media_format = "h264/aac" metadata =
"main_view/spanish" temi_init = "3699255471291907022">
<source protocol = "http/dash" uri = "http://158.42.X.X/multicam/dash/cam1.mpd"/>
</MEDIA>

<MEDIA id = "media2" media_type = "AV" media_format = "h264/aac" metadata =
"left_view/spanish" temi_init = "3699255471291907022">
<source protocol = "http/dash" uri = "http://158.42.X.X/multicam/dash/cam2.mpd"/>
</MEDIA>

<MEDIA id = "media3" media_type = "AV" media_format = "h264/aac" metadata =
"alternative_front_view/spanish" temi_init = "3699255471291907022">
<source protocol = "http/dash" uri = "http://158.42.X.X/multicam/dash/cam3.mpd"/>
</MEDIA>

<MEDIA id = "media4" media_type = "AV" media_format = "h264/aac" metadata =
"right_view/spanish" temi_init = "3699255471291907022">
<source protocol = "http/dash" uri = "http://158.42.X.X/multicam/dash/cam4.mpd"/>
</MEDIA>

<IDMS id = "idms_mgr" protocol = "websocket" metadata = "idms manager uri SMS" uri =
"http:// 158.42.X.X:YY/idms_manager1"/>

<CLOCK id = "global_clock" protocol = "ntp" media_type = "time" media_format =
"64_bit_ntp_time" metadata = "" uri = "158.42.X.X "/>
```

```

<LASTUPDATE protocol = "http" media_type = "time" metadata = "" format = "mm/dd/yyyy-
hh:mm:ss" value = "31/10/2018-23:59:59"/>
</ Related Media Contents File >

```

Fig. 87 Fichero RMCF utilizado en el escenario IDMS

6.4.2 METODOLOGÍA DE EVALUACIÓN EMPLEADA

En este caso, la solución IDMS implementada en el escenario se ha evaluado a partir de los valores de asincronía registrados por el dispositivo principal que es el que ejecuta la MS app, y, por tanto, reproduce el contenido broadcast (principal) en cada hogar (o destinatario). Tal y como se aprecia en la Fig. 86, también se ha utilizado contenido de vídeo con el número de fotograma superpuesto, con el fin de poder realizar una primera evaluación visual más rápida y poder comprobar así el funcionamiento del mecanismo de sincronización.

Para la evaluación se han diseñado 8 casos diferentes, combinando las siguientes condiciones: i) adopción de un esquema de control de sincronización centralizado (SMS) o un esquema de control de sincronización distribuido (DCS); ii) condiciones ideales o con pérdida de paquetes en la transmisión broadcast (probabilidad de pérdida configurada a 0.001%, valor significativamente mayor que el valor medio en QEF) y retardo en la red broadband (retardo configurado con un valor de $60\text{ms} \pm 139$

siguiendo una distribución normal); y iii) referencia de sincronización al dispositivo principal (MS app) con el punto de reproducción más atrasado o a un punto promedio de la reproducción de todos los dispositivos principales (apps MS). En esta evaluación, no se ha limitado el ancho de banda disponible a través de la red broadband, puesto que la gestión del cambio de una calidad a otra por parte del proceso de reproducción de contenido se ha validado satisfactoriamente en los escenarios IDES anteriormente descritos. La Tabla 24 resume todos los casos definidos y se indica en cada uno de ellos de qué solución de sincronización (IDMS o IDES) se han registrado asincronías.

Tabla 24 Resumen de los casos considerados y tipo(s) de sincronización evaluado(s) en cada uno de ellos

Referencia de sincronización	SMS		DCS	
	Condiciones ideales	Condiciones no-ideales	Condiciones ideales	Condiciones no-ideales
Al punto de reproducción de la MS app más atrasada	IDMS	IDMS, IDES	IDMS	IDMS, IDES
Punto de reproducción promedio	IDMS	IDMS, IDES	IDMS	IDMS, IDES

Tal y como se puede observar, para los casos en los que la condiciones son ideales únicamente se registran los valores de asincronías entre los procesos de reproducción de las MS apps en cada destinatario (IDMS) y, por el contrario, cuando las condiciones no son ideales y existe una probabilidad de pérdida de paquetes en la transmisión broadcast y retardos en la red broadband, también se registran los niveles de asincronías entre el proceso de reproducción de la MS app en el dispositivo principal y el proceso de reproducción de las CS apps en los dispositivos secundarios locales (IDES) de cada hogar. Esto permite valorar en qué magnitud afectan los ajustes (debidos a la solución de sincronización IDMS) en el proceso de reproducción de la MS app y los errores provocados, sobre todo por la pérdida de paquetes en la transmisión broadcast en el proceso de reproducción de la CS app en los dispositivos secundarios.

6.4.3 RESULTADOS OBTENIDOS

A continuación, se muestran los resultados obtenidos para cada uno de los 8 casos propuestos. Para cada uno de ellos se han realizado 10 sesiones de aproximadamente 3 minutos cada una,

con el fin de que los dispositivos principales y secundarios involucrados hayan podido alcanzar un estado de sincronización válido y un proceso de reproducción estable.

ESQUEMA DE CONTROL CENTRALIZADO SYNCHRONIZATION MAESTRO SCHEME

En esta sección se exponen los resultados obtenidos adoptando un esquema de control SMS para el mecanismo de sincronización IDMS. En la Tabla 25 se muestran los resultados obtenidos según las condiciones de las redes de transmisión (es decir, ideal o con pérdidas de paquetes en la transmisión broadcast y retardo de red); y según la referencia seleccionada (al punto de reproducción de la MS app más atrasado o a un punto de reproducción promedio), a partir de los valores registrados en cada uno de los 4 dispositivos principales (ejecutando la MS app) involucrados. En la figura se representa el RMS medio de los valores de la asincronía de cada una de las 10 sesiones (*session async*, definida en la sección 4.6.5) registrados en cada caso, junto con el I.C. 95%.

Tabla 25 Resultados obtenidos del RMS medio e I.C. 95% de los valores de la asincronía registrados en las 10 sesiones realizadas en cada caso respecto a la solución de sincronización IDMS con un esquema de control SMS

Condiciones	Más atrasado (ms)	Punto promedio (ms)
Ideal	150.07 ± 2.55	150.91 ± 2.23
Pérdidas + retardos	157.11 ± 3.73	153.37 ± 5.87

Los valores presentados en la Tabla 25 se han calculado a partir de los valores obtenidos en los 4 dispositivos principales (ejecutando la MS app) involucrados. Se puede observar que el valor RMS de los valores de la asincronía registrada obtenido no llega a superar el umbral *session async*, configurado en 160ms. Además, se puede observar cómo los valores de RMS de los valores de la asincronía registrada obtenidos son mayores cuando las condiciones no son ideales, es decir, cuando existe una cierta probabilidad de pérdida de paquetes en la transmisión broadcast y un cierto retardo variable en la red broadband. La Fig. 88 presenta el número medio de ajustes que se han necesitado realizar en el proceso de reproducción de la MS app para mantener un nivel de asincronía IDMS por debajo del umbral configurado de ±160ms. Se trata de una información relevante, puesto que el número de ajustes y su magnitud pueden afectar a la QoE percibida por el consumidor. Cuantos menos ajustes y de menor magnitud se realicen en los procesos de reproducción, mejor será la QoE de los usuarios.



Fig. 88 Número medio de ajustes necesarios durante la evaluación IDMS - SMS con 4 destinatarios involucrados

Se puede observar que, cuando la referencia de sincronización es el dispositivo cuyo proceso de reproducción de la MS app es el más atrasado (en cuanto al instante de reproducción del contenido), el número medio de ajustes que se realiza durante una sesión es ligeramente mayor que cuando la referencia es un punto de reproducción promedio entre las MS apps de los 4 dispositivos principales involucrados. Además, se aprecia que las diferencias existentes cuando un

escenario presenta condiciones ideales frente al mismo escenario con pérdidas de paquetes en la transmisión broadcast y retardos en la red broadband no son significativas. Esto se puede deber a que el posible retardo provocado en el proceso de reproducción por la pérdida de paquetes sea de una magnitud menor que la magnitud de los ajustes en el proceso de reproducción del contenido broadcast (en la MS app) debidos a la solución de sincronización IDMS, y, por tanto, poco significativo.

ESQUEMA DE CONTROL DISTRIBUIDO *DISTRIBUTED CONTROL SCHEME*

En esta sección se exponen los resultados obtenidos adoptando un esquema de control de sincronización distribuido o DCS para la solución de sincronización IDMS. En la Tabla 26 se muestran los resultados obtenidos según las condiciones de las redes de transmisión (es decir, ideal o con pérdidas de paquetes en la transmisión broadcast y retardo de red); y según la referencia seleccionada (al punto de reproducción de la MS app más atrasado o a un punto de reproducción promedio), a partir de los valores registrados en cada uno de las 4 MS app involucradas.

Tabla 26 Resultados obtenidos del RMS medio e I.C. 95% de los valores de la asincronía registrados en las 10 sesiones realizadas en cada caso respecto a la solución de sincronización IDMS con un esquema de control DCS

Condiciones	Más atrasado (ms)	Punto promedio (ms)
Ideal	148.61 ± 3.08	133.68 ± 13.58
Pérdidas + retardos	155.80 ± 8.42	153.08 ± 5.96

Los valores de la Tabla 26 son resultado del promedio entre los valores registrados en las 10 sesiones realizadas por las 4 MS apps involucradas. Se puede comprobar que los valores registrados tampoco exceden el umbral *session async* configurado. Sin embargo, para el caso en el que la referencia es el punto de reproducción de la MS app más atrasada y las condiciones involucran pérdidas de paquetes en la transmisión broadcast y retardos en la red broadband, el valor resultante de sumar el nivel RMS promedio alcanzado más el I.C. 95% sí es superior al umbral establecido. Esto se puede explicar debido a la aleatoriedad de la pérdida de paquetes en la red broadcast y del valor de retardo de la red broadband emulado, puesto que estos factores afectan negativamente a los procesos de reproducción involucrados y cada MS app puede verse afectada de manera distinta e independientemente del funcionamiento del resto de apps MS involucradas. La Fig. 89 muestra los valores registrados por cada MS app, lo que permite entender mejor el motivo por el cual se llega a exceder el umbral configurado en el caso mencionado anteriormente. Se observa cómo el dispositivo ejecutando la MS app1 es el que provoca ese aumento en el valor promedio de la asincronía de la sesión alcanzada cuando las condiciones no son ideales y la referencia es el punto de reproducción de la MS app más atrasada.

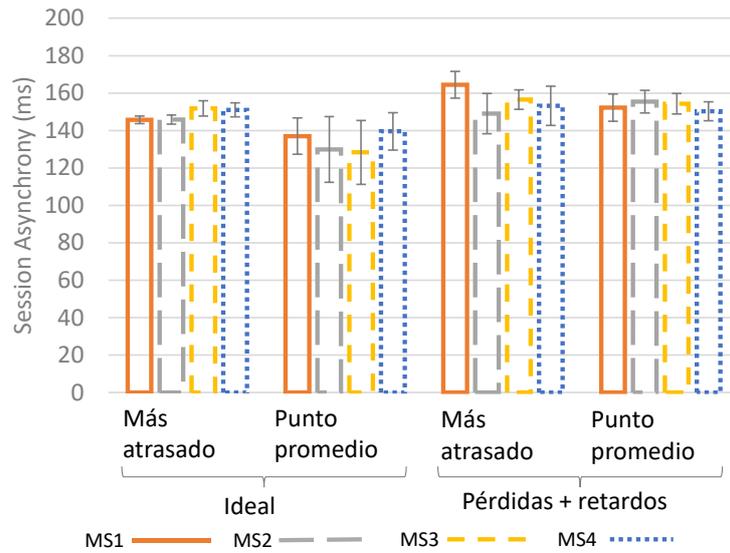


Fig. 89 Valor medio del RMS e I.C. 95% de los valores de la asincronía registrados en las 10 sesiones realizadas, obtenidos en cada MS app y en cada caso, respecto a la solución de sincronización IDMS con un esquema de control DCS

Adicionalmente, en la Fig. 90 se muestra el número medio de ajustes realizados en cada caso.

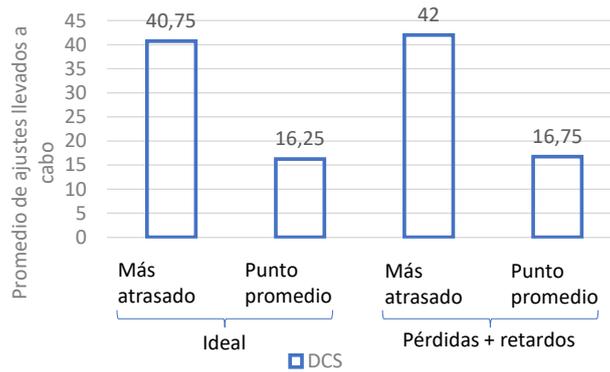


Fig. 90 Número medio de ajustes realizados durante la evaluación IDMS - DCS con 4 destinatarios involucrados

A continuación, en la Fig. 91, se muestra una comparativa del número de ajustes requeridos por las MS app dependiendo el esquema de control IDMS adoptado.

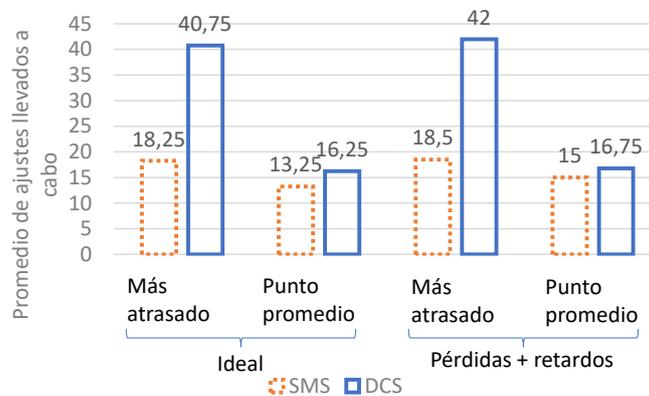


Fig. 91 Número medio de ajustes realizados durante la evaluación IDMS - DCS con 4 destinatarios involucrados

Si se comparan los resultados del número de ajustes necesarios en el proceso de reproducción de la MS app bajo un esquema SMS o DCS, se puede comprobar cómo el esquema de control DCS requiere que las MS apps realicen en promedio un mayor número de ajustes para mantener el nivel de sincronización por debajo del umbral *session async* establecido. Para los dos casos analizados (con condiciones ideales o con pérdidas de paquetes en la transmisión broadcast y retardos en la red broadband) en los que la referencia es un punto de reproducción promedio, la diferencia entre el número de ajustes realizados del esquema de control DCS respecto al esquema de control SMS no es significativo, aunque el número medio de ajustes sea ligeramente superior. Sin embargo, cuando la referencia pasa a ser el dispositivo con el punto de reproducción de la MS app más atrasada, el número medio de ajustes necesarios cuando se adopta el esquema de control DCS es más del doble que los necesarios si se adopta un esquema de control SMS.

RENDIMIENTO DEL MECANISMO DE SINCRONIZACIÓN HÍBRIDA INTER-DISPOSITIVO EN ESCENARIOS INTER-DESTINATARIO

En esta sección se presentan los resultados del rendimiento del mecanismo IDES en el escenario IDMS, en el que las MS apps de los dispositivos principales involucrados en la sesión pueden sufrir ajustes en sus procesos de reproducción y, de forma indirecta, afectar a los procesos de reproducción de los contenidos secundarios (es decir, de las CS apps) debido al propio mecanismo de sincronización IDES implementado. Para esta evaluación particular, se han considerado únicamente escenarios con pérdidas de paquetes en la transmisión broadcast y retardos en la red broadband. Tal y como se expone en la Tabla 23, los dispositivos utilizados como CS apps tienen prestaciones y especificaciones heterogéneas, por lo que los resultados obtenidos varían según el dispositivo. La Tabla 27 y la Tabla 28 muestran los valores medios de las asincronías obtenidos en la sincronización híbrida IDES alcanzada, junto con el I.C. 95%, para los dos esquemas de control, SMS y DCS, respectivamente.

Tabla 27 Valor cuadrático medio e I.C. 95% de los valores de las asincronías registradas con un esquema de control centralizado SMS

Dispositivo	Referencia al más atrasado	Referencia a un punto de reproducción promedio
S5	77.473ms ± 3.820ms	70.424ms ± 6.164ms
S6	76.064ms ± 7.820ms	57.925ms ± 8.245ms
Tab 1	74.176ms ± 9.470ms	65.087ms ± 11.700ms
Tab 2	72.945ms ± 10.975ms	59.153ms ± 8.490ms

Tabla 28 Valor cuadrático medio e I.C.95% de los valores de las asincronías registradas con un esquema de control distribuido DCS

Dispositivo	Referencia al más atrasado	Referencia a un punto de reproducción promedio
S5	78.121ms ± 3.585ms	75.284ms ± 5.920ms
S6	77.823ms ± 7.965ms	75.786ms ± 6.138ms
Tab 1	79.316ms ± 5.130ms	75.464ms ± 5.191ms
Tab 2	79.775ms ± 4.975ms	75.453ms ± 8.170ms

En la Fig. 92 se muestran los valores de las tablas anteriores de forma gráfica. Se aprecian mejor las diferencias del funcionamiento del mecanismo de sincronización híbrida IDES según el esquema de control y referencia de sincronización adoptados.

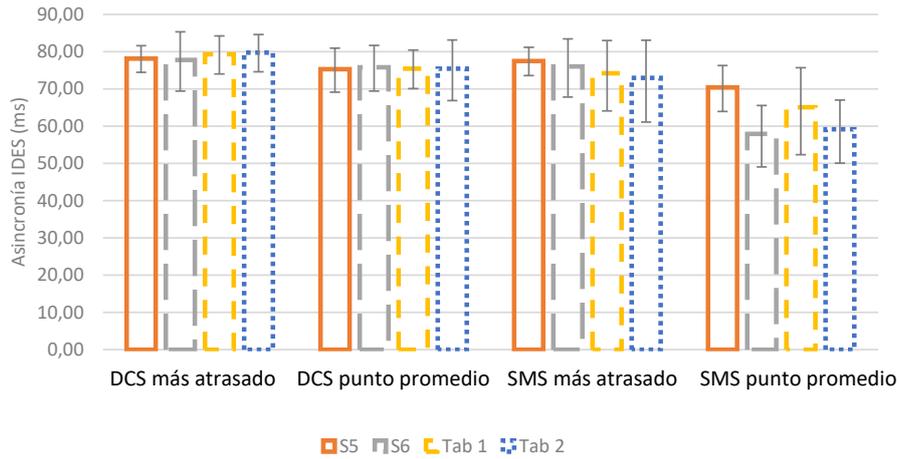


Fig. 92 Valor medio del RMS e I.C. 95% de los valores de la asincronía registrados en las 10 sesiones realizadas, obtenidos en cada CS app durante la evaluación IDMS

Según los resultados obtenidos, el valor RMS de los valores de asincronía registrados por las CS apps no supera el umbral IDES configurado de 80ms. Sin embargo, si se consideran los valores obtenidos del I.C. 95%, en algunos casos dicho umbral sí se supera ligeramente. Esto se puede explicar debido a que, durante una sesión IDMS, la MS app realiza ajustes en su propio proceso de reproducción del contenido principal debido a la solución de sincronización IDMS y dichos ajustes producirán asincronías con respecto a los procesos reproductores de las CS apps (de los contenidos complementarios) que, si superan un cierto umbral, también deberán realizar ajustes forzados por la solución de sincronización IDMS, para seguir sincronizados con la reproducción del contenido principal. Es por ello por lo que se produce un aumento del número de ajustes y de su magnitud en el proceso de reproducción de las CS apps. La Fig. 93 muestra, para cada caso y dispositivo, la distribución de frecuencia acumulada. A través de esta figura, se puede observar con mayor claridad qué dispositivos presenta un mejor rendimiento respecto al nivel de sincronización alcanzado, puesto que muestra en probabilidades (tanto por uno, eje de coordenadas Y) la aparición de un determinado valor de sincronización (eje de coordenadas X).

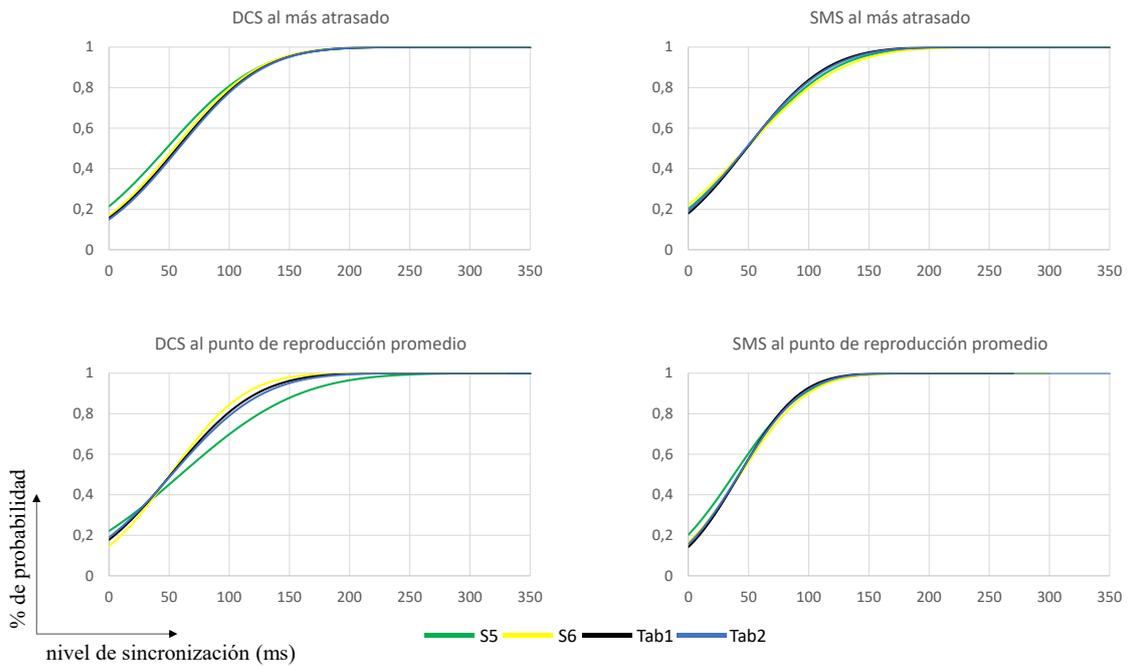


Fig. 93 Función de frecuencia acumulada de los valores absolutos de las asincronías registradas durante la evaluación objetiva en cada dispositivo secundario para cada uno de los escenarios IDMS evaluados (con esquemas de control SMS y DCS)

Como se puede observar, en términos generales, el dispositivo secundario smartphone Samsung Galaxy S5 también ofrece el peor rendimiento en cuanto al nivel de sincronización obtenida en este caso. Además de ofrecer unos valores del RMS medio de los valores de asincronía registradas más próximos al umbral configurado, en la Fig. 93 se puede ver cómo su comportamiento difiere de todos los demás dispositivos secundarios que convergen de manera parecida. Este rendimiento más bajo se puede explicar por las peores prestaciones del dispositivo, ya que se trata del más antiguo utilizado durante las evaluaciones.

Adicionalmente, se puede deducir que cuando se utiliza la solución de sincronización IDMS con un esquema de control centralizado SMS y se toma como referencia un punto de reproducción promedio de todas las apps MS, es el caso en el que los dispositivos secundarios alcanzan un mejor rendimiento en cuanto al nivel de sincronización híbrida IDEs. La Tabla 29 aporta, en base a lo expuesto en la Fig. 93, las probabilidades (promediadas entre los cuatro dispositivos secundarios) de registrar asincronías entre la reproducción del contenido principal y los contenidos secundarios, con valores absolutos por debajo de 80ms (umbral IDEs configurado).

Tabla 29 Probabilidad de registrar un valor de asincronía (valor absoluto) de hasta 80ms

Esquema / Referencia	Probabilidad
SMS / Más lento	72%
SMS / Punto de reproducción promedio	82%
DCS / Más lento	65%
DCS / Punto de reproducción promedio	68%

Finalmente, se han realizado cuatro sesiones adicionales (una para cada caso según el esquema de control de sincronización y referencia adoptado, incluyendo pérdida de paquetes en la transmisión broadcast y retardos en la red broadband), con una duración aproximada de 5 minutos, para evaluar el comportamiento y los valores de asincronías registrados por cada dispositivo secundario cuando durante las sesiones se realizan cambios en los contenidos

secundarios a ser visualizados. En cada caso, uno o dos dispositivos secundarios realizan dichos cambios. De la Fig. 94 a la Fig. 97 se aprecia el comportamiento de los diferentes dispositivos secundarios involucrados. Se puede comprobar cómo el dispositivo secundario smartphone Samsung Galaxy S5 realiza un número significativamente mayor de ajustes que el resto de los dispositivos secundarios (con mejores prestaciones técnicas) que ofrecen una mayor estabilidad en el proceso de reproducción con un menor número de ajustes.

Al igual que en la evaluación de la plataforma en un escenario IDES, los dispositivos secundarios (sus CS apps) reproducen contenido DASH adquirido a través de la red broadband. Cada vez que se realiza un cambio de contenido, el nivel de sincronización se degrada considerablemente. Sin embargo, se puede comprobar en las gráficas cómo rápidamente se realizan los primeros ajustes de sincronización hasta que se alcanza un nivel de sincronización híbrida IDES estable.

En la Fig. 94 se observan los valores de asincronía IDES registrados en la sesión en la que se adoptó un esquema de control de la sincronización centralizado SMS y una referencia de sincronización el punto de reproducción de la MS app del dispositivo principal más atrasado. En este caso, el único dispositivo que realizó un cambio en el contenido es el smartphone Samsung Galaxy S6, en el segundo 160, aproximadamente. De nuevo, el smartphone Samsung Galaxy S5 es el dispositivo que peor rendimiento ofrece. Si, más adelante, se compara el comportamiento de los dispositivos secundarios cuando el esquema de control de la sincronización es distribuido DCS y la referencia de sincronización es el punto de reproducción de la MS app del dispositivo principal más atrasado (mostrado más adelante en la Fig. 96), se observa que el número de ajustes necesarios es menor si el esquema de control es SMS.

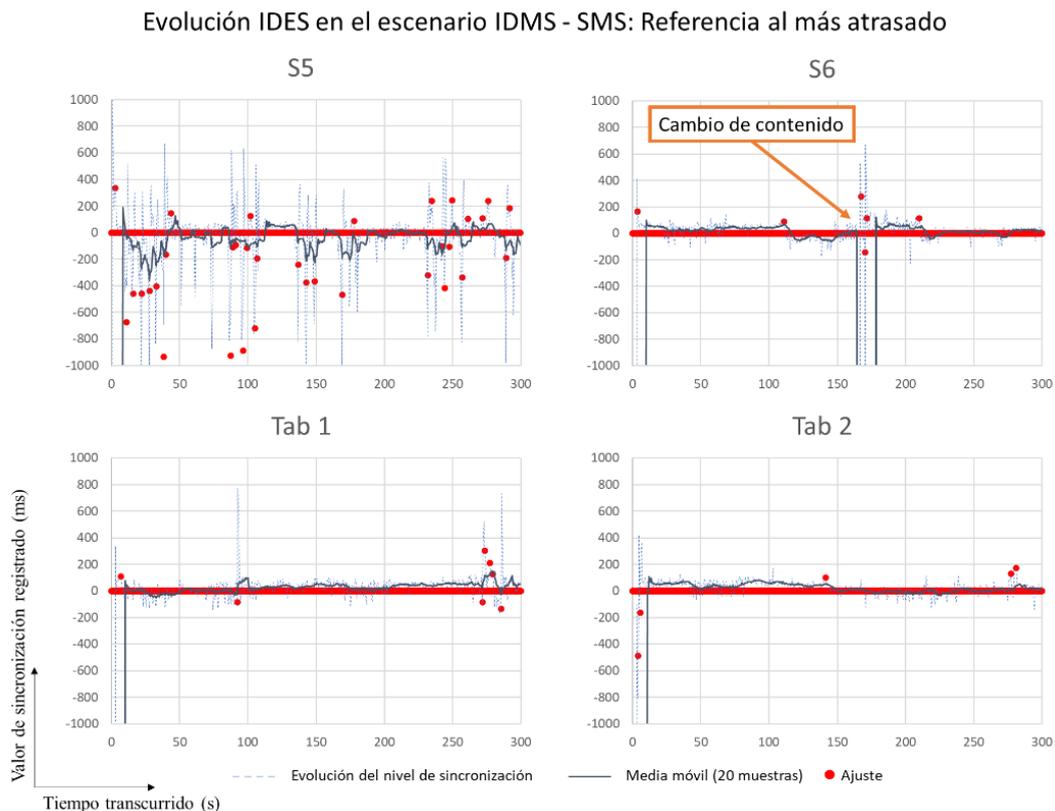


Fig. 94 Comportamiento de los dispositivos secundarios en un escenario IDMS – SMS con referencia de sincronización el punto de reproducción de la MS app del dispositivo principal más atrasado

En la Fig. 95 se muestra el comportamiento de los dispositivos secundarios en la sesión en la que se adoptó el esquema de control de la sincronización SMS y la referencia el punto promedio

de la reproducción de las apps MS de los 4 dispositivos principales involucrados. En este caso, realizaron cambios del contenido visualizado los dispositivos secundarios smartphone Samsung Galaxy S6, en los segundos 25 y 175, y la Tab 2 (Samsung Galaxy Tab S), en el segundo 200, aproximadamente. Una vez más, el dispositivo secundario smartphone Samsung Galaxy S5 es el dispositivo que ofrece el peor rendimiento, a pesar de que, en este caso, dicho dispositivo no ha cambiado de contenido en ningún momento. Si se compara el comportamiento de los dispositivos secundarios cuando el esquema de control de la sincronización adoptado es el centralizado SMS, pero la referencia es el más atrasado (Fig. 94), se puede observar que el rendimiento que ofrecen los dispositivos secundarios mejora ligeramente, puesto que el número de ajustes necesarios es menor y la estabilidad en los procesos de reproducción de los mismos es mejor, aunque la diferencia en el número de ajustes no se puede considerar significativa.

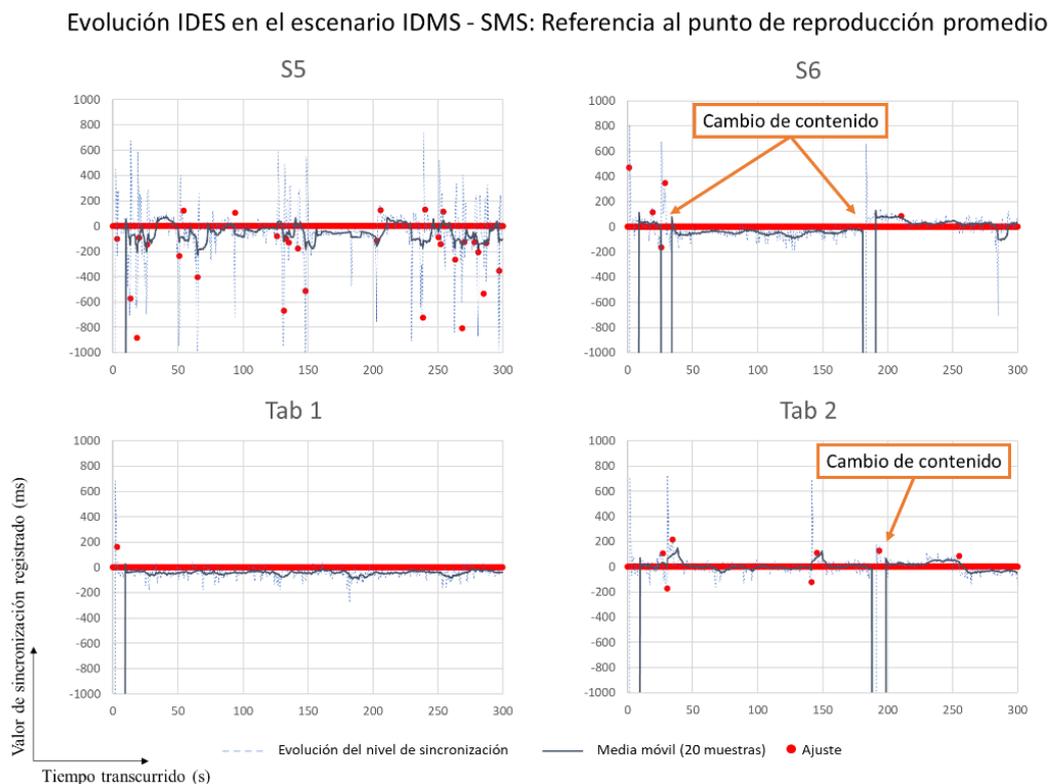


Fig. 95 Comportamiento de los CS en un escenario IDMS – SMS con referencia de sincronización el punto promedio de la reproducción de las apps MS de los 4 dispositivos principales involucrados

La Fig. 96 muestra el comportamiento de los dispositivos secundarios en la sesión en la que se adoptó un esquema de control de sincronización distribuido o DCS, con la referencia el punto de reproducción de la MS app del dispositivo principal más atrasado. Alrededor del segundo 140, el dispositivo Tab 2 (Samsung Galaxy Tab S) realiza un cambio del contenido complementario a reproducir en él. Se puede observar cómo los dispositivos secundarios Samsung Galaxy S6 y las dos tablets Samsung Galaxy Tab S ofrecen un rendimiento similar en cuanto a la frecuencia de ajustes y valores de asincronías registrados. Sin embargo, se puede comprobar que el dispositivo secundario Samsung Galaxy S5 realiza más ajustes en su proceso de reproducción, a pesar de estar ejecutando la misma aplicación CS que el resto de los dispositivos. Al realizar más ajustes, el rendimiento y la estabilidad en el proceso de reproducción empeoran.

Evolución IDES en el escenario IDMS - DCS: Referencia al más atrasado

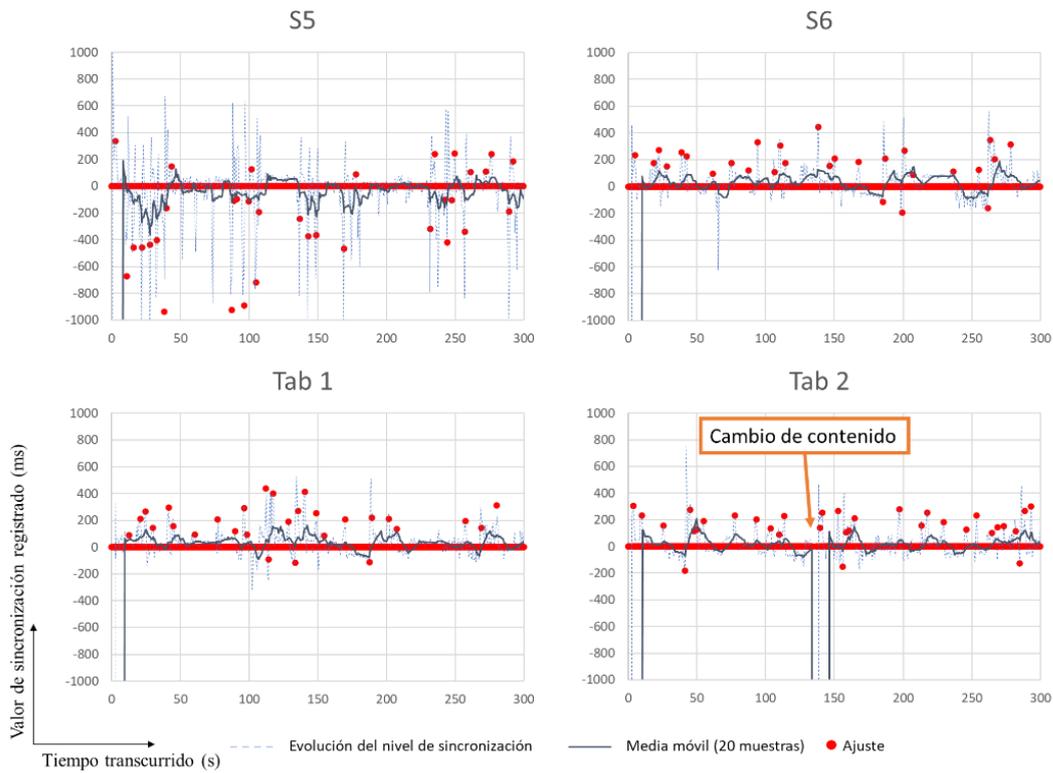


Fig. 96 Comportamiento de los dispositivos secundarios en un escenario IDMS – DCS con referencia de sincronización el punto de reproducción de la MS app del dispositivo principal más atrasado

La Fig. 97 muestra el comportamiento de los dispositivos secundarios en la sesión en la que se adoptó un esquema de control de sincronización distribuido DCS y como referencia de sincronización un punto promedio de la reproducción de las apps MS de los 4 dispositivos principales involucrados. En este caso, el dispositivo secundario Samsung Galaxy S5 realiza un cambio de contenido complementario aproximadamente en el segundo 200, mientras que el dispositivo secundario Samsung Galaxy S6 lo realiza en los segundos 110, 130 y 200, aproximadamente. De forma similar a la Fig. 96, en este caso es también el dispositivo secundario Samsung Galaxy S5 el que ofrece con un peor rendimiento respecto al nivel de sincronización alcanzado y al número de ajustes realizados en el proceso de reproducción. Por otro lado, comparando este caso con el caso anterior en el que la referencia de sincronización era el punto de reproducción de la MS app más atrasada (manteniendo el esquema de control DCS), se puede comprobar que en este caso el número de ajustes necesarios para mantener los valores de asincronía por debajo del umbral IDES configurado es menor.

Evolución IDES en el escenario IDMS - DCS: Referencia al punto de reproducción promedio

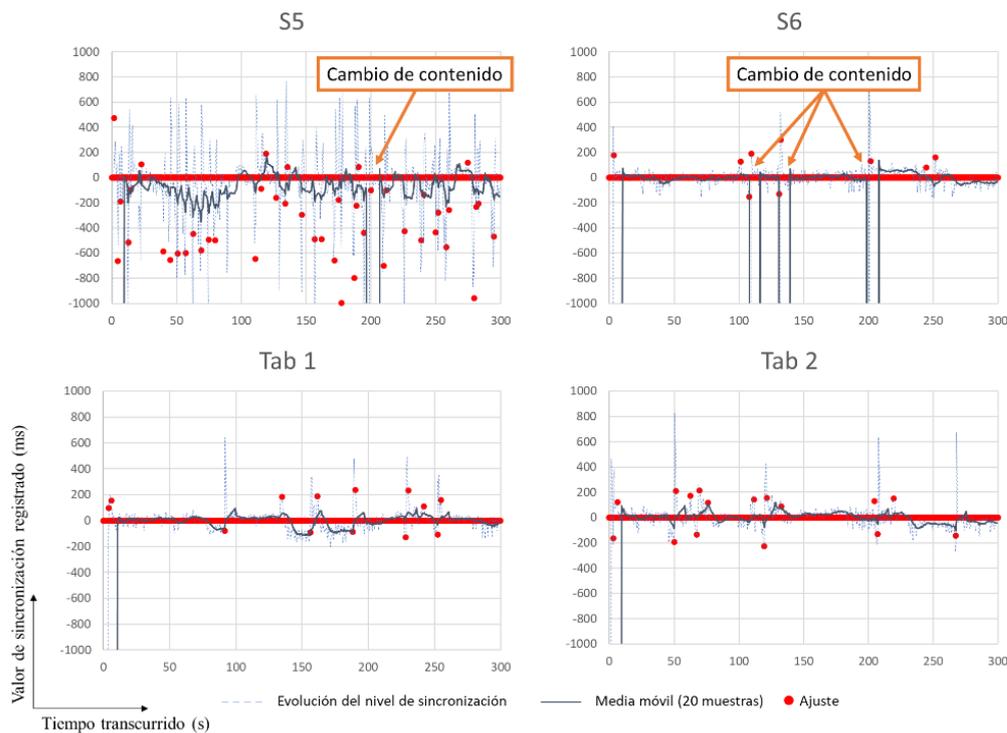


Fig. 97 Comportamiento de los dispositivos secundarios en un escenario IDMS – DCS con referencia de sincronización el punto promedio de la reproducción de las apps MS de los 4 dispositivos principales involucrados

6.4.4 DISCUSIÓN

En la evaluación del escenario IDMS se ha podido comprobar que un esquema de control centralizado SMS ofrece un mejor rendimiento respecto al nivel de sincronización IDMS alcanzado, además de necesitar un menor número de ajustes en el proceso de reproducción para alcanzar valores de asincronía por debajo del umbral *session async* configurado. Sin embargo, tal y como se explica en capítulos anteriores, si el dispositivo encargado de ejecutar el módulo IDMS Manager sufre algún error o desconexión, el nivel de sincronización puede verse comprometido. Este es el punto favorable respecto al esquema de control distribuido DCS, puesto que los valores de asincronía también se mantienen por debajo del umbral *session async* configurado y este tipo de esquema ofrece una mayor tolerancia a fallos.

Respecto al rendimiento del mecanismo IDES bajo un escenario IDMS, se ha comprobado que el nivel de sincronización híbrida IDES alcanzado se ve claramente afectado, sobre todo, cuando el esquema de control adoptado es el distribuido DCS. Esto es debido a que, al tratarse de un esquema M/S el adoptado para la sincronización IDES entre los procesos de reproducción de las MS y CS apps, conforme la MS app realiza algún ajuste en el proceso de reproducción del contenido principal, éste afecta directamente al proceso de reproducción del contenido complementario reproducido en la CS app. A los ajustes en el proceso de sincronización híbrida IDES ya necesarios de por sí, se le añaden los provocados por ajustes en el proceso de reproducción de la MS app debidos al mecanismo de sincronización IDMS. Sin embargo, tras observar los resultados presentados, se pueden considerar que la combinación de ambos mecanismos de sincronización IDES e IDMS producen resultados satisfactorios. Como promedio, no se superan los respectivos umbrales de asincronía configurados. Salvo situaciones excepcionales (por ejemplo, un error inducido por una pérdida de paquete en la transmisión broadcast que afecte a la reproducción del MS o al envío de la información temporal necesaria a la CS app), las MS y CS apps presentan una reproducción sincronizada y estable de los contenidos

híbridos tanto en escenarios IDES, como en escenarios IDMS (incluyendo contenidos complementarios que requieren de mecanismos de sincronización híbrida IDES). Si bien, es cierto que, en términos generales, uno de los factores más significativos por el que esperar un mejor o peor funcionamiento de los mecanismos de sincronización es las especificaciones y capacidades de los dispositivos involucrados, tal y como se ha podido observar a partir del rendimiento ofrecido por el dispositivo secundario de peores prestaciones, el smartphone Samsung Galaxy S5.

6.5 CONCLUSIONES

En este capítulo se han presentado los escenarios, la metodología seguida y los resultados de las evaluaciones objetivas llevadas a cabo durante la Tesis Doctoral, con respecto al nivel de sincronización híbrida IDES e IDMS alcanzado. Se ha evaluado la sincronización en varios escenarios: un primer escenario básico incluyendo una solución IDES (sin contenido omnidireccional), el cual incluye contenido multivista; un segundo escenario inmersivo incluyendo una solución IDES con contenido omnidireccional; y finalmente, un tercer escenario incluyendo soluciones de sincronización híbrida IDES e IDMS.

En los escenarios de evaluación IDES, se ha demostrado que, aun habiendo pérdidas de paquetes en la transmisión broadcast, y retardos en la transmisión broadband, el nivel de sincronización entre las MS y CS apps se mantienen siempre por debajo del umbral de asincronía máxima configurado.

Por otro lado, en el escenario IDMS también se ha corroborado que los valores de asincronía registrados entre destinatarios se mantienen por debajo del umbral de la asincronía de la sesión máxima configurado. Respecto a la combinación de soluciones de sincronización IDMS + IDES, se observa cómo los ajustes IDMS en una MS app afectan de forma directa a los ajustes de la CS app enlazada, pudiendo afectar negativamente en la QoE percibida por los usuarios en momentos puntuales.

Cabe resaltar que, tras los resultados obtenidos en los diferentes escenarios presentados, contar con un dispositivo secundario con prestaciones medias/altas puede suponer una mejora significativa en la QoE percibida, ya que un dispositivo con recursos insuficientes puede provocar una mala gestión de la recepción y presentación del contenido, influyendo así de manera directa en el número de ajustes requeridos para alcanzar un nivel de sincronización aceptable.

Capítulo 7. EVALUACIÓN SUBJETIVA Y RESULTADOS

En este capítulo se presentan las diferentes evaluaciones subjetivas que se han llevado a cabo, su metodología, así como los resultados obtenidos y su discusión.

7.1 INTRODUCCIÓN

El éxito de una plataforma o una aplicación multimedia radica principalmente en la aceptación de la misma por parte de los usuarios finales. Esta aceptación depende de diversos aspectos como, por ejemplo, la facilidad de uso (esto es, la usabilidad), su interfaz gráfica de usuario, el rendimiento de los procesos involucrados, la utilidad o ventajas de sus funcionalidades y su aplicabilidad en escenarios de interés para el usuario final.

Por tanto, con el fin de determinar el éxito o el buen funcionamiento de los casos de uso implementados en la plataforma desarrollada incluyendo los mecanismos de sincronización IDES e IDMS propuestos y efectos multisensoriales, se debe tener en cuenta el grado de aceptación de los consumidores o usuarios finales. En primer lugar, han sido evaluados por usuarios casos de uso implementados en la plataforma en los tres escenarios descritos en el capítulo anterior (escenario IDES básico con contenido multivista, un escenario IDES inmersivo con contenido omnidireccional disponible; un escenario de TV Social, incluyendo una solución de sincronización IDMS), además de un cuarto un escenario IDES mulsemmedia, incluyendo contenido 2D, omnidireccional y efectos multisensoriales.

En primer lugar, se presentan una serie de directrices seguidas en todas las evaluaciones. A continuación, dentro de la sección de evaluación de cada escenario (IDES básico, IDES inmersivo, IDMS y Mulsemmedia) se presenta el perfil de los participantes, la metodología de evaluación empleada y los resultados obtenidos.

7.2 DIRECTRICES COMUNES EN TODAS LAS EVALUACIONES REALIZADAS POR LOS PARTICIPANTES

Todas las evaluaciones subjetivas llevadas a cabo para la validación de las soluciones de sincronización diseñadas y propuestas en esta Tesis Doctoral, han tenido en común los siguientes puntos:

- El proceso de evaluación ha seguido las directrices definidas en la recomendación ITU-R BT.500-13 [166], con el objetivo de preparar el entorno que emula el hogar adecuadamente, y que el proceso no dure más de 30 minutos.
- Los participantes han realizado una lectura previa de un documento breve con la descripción del escenario y el caso de uso a evaluar, así como la cumplimentación de un

cuestionario inicial (en adelante, fase 0). Previamente a cualquier interacción con el escenario, los participantes han rellenado un formulario donde se recoge información personal, como, por ejemplo, el género o la edad, el perfil profesional, etc.

- En primer lugar, los participantes han realizado un uso guiado de las aplicaciones de la plataforma para el escenario objeto de evaluación (en adelante, fase 1). Con el fin de que los participantes se puedan familiarizar con el escenario y sus funcionalidades, han sido monitorizados por un revisor durante unos minutos con el fin de resolver dudas y mostrar el funcionamiento y características del escenario objeto de evaluación.
- En segundo lugar, los participantes han realizado un uso libre de la plataforma (en adelante, fase 2). Una vez los participantes se han familiarizado con la plataforma y el escenario objeto de evaluación, han tenido un periodo de tiempo de uso libre de las aplicaciones y de las funcionalidades disponibles en la plataforma para cada escenario.
- En tercer lugar, los participantes han cumplimentado un cuestionario final (en adelante, fase 3) para obtener su opinión acerca de la experiencia global de consumo y la usabilidad del escenario implementado (si corresponde) en la plataforma desarrollada en la Tesis Doctoral. Para dicho fin se ha utilizado la escala SUS o *System Usability Scale*, definida en [167]. Además, se valora el nivel de sincronización percibido entre el dispositivo principal (MS app) y los diferentes dispositivos secundarios (CS apps) que puedan existir para cada tipo de escenario, empleando la metodología más común para medir la QoE, que actualmente se basa en el análisis de la experiencia percibida por el usuario a través de la métrica MOS (Mean Opinion Score, ITU-T P.800 [168]). Para ello se han utilizado cuestionarios post experiencia multisensorial con escalas de valoración de tipo Likert de cinco niveles.

7.3 ESCENARIO INTER-DISPOSITIVO BÁSICO

En esta sección se describe el proceso de evaluación subjetiva del caso de uso implementado en un escenario IDES básico, consistente en el caso de uso de contenidos híbridos multivista, en el que el contenido complementario disponible es contenido de múltiples vistas de la escena representada en el contenido principal (broadcast) y se podrá visualizar en múltiples dispositivos. En dicho escenario, entre otras cosas, se evaluará el grado de precepción por parte de los usuarios participantes en el estudio de la sincronización obtenida mediante el mecanismo de sincronización IDES integrado en la plataforma para dicho caso de uso. Respecto al contenido y dispositivos utilizados en esta evaluación son los mismos que se emplearon para la evaluación objetiva del escenario IDES básico (ver sección 6.2) y que se adjuntan en la Tabla 30 y Tabla 31.

Tabla 30 Contenidos híbridos utilizados en el escenario IDES básico

Technologie	Contenido	Codificación
Broadcast (MPEG2-TS)	Vídeo	H.264, 1920x1080, 25fps (4 canales, 30Mbps)
	Audio	MPEG-4 AAC 140Kbps (2 canales, 48KHz)
Broadband (DASH)	Perfil de vídeo	H.264, 25 fps, duración de los segmentos: 3s
		Calidad 1 (Q1): 426x240, 700Kbps
		Q2: 640x360, 1000Kbps
		Q3: 854x480, 2000Kbps
		Q4: 1280x720, 4000Kbps
		Q5: 1920x1080, 6000Kbps
	Audio	MPEG-4 AAC 128Kbps (2 canales, 48KHz)

Tabla 31 Dispositivos utilizados en el escenario IDES básico

Dispositivo	Características
PC (broadcaster)	Intel Core i7-6700 @ 3.40GHz, 8GB RAM, SSD 240GB, Windows 10, DVB-T PCI Card: DekTec DTA-2111
PC (servidor multimedia y NTP)	Intel Xeon E5420 @ 2.50GHz x8, 8GB RAM, HDD 200GB, Ubuntu 14.04, Apache Server v2.4.7, Fast Ethernet embedded card.
Switch/Router	TP-Link AC1900 Wireless Dual Band Gigabit Router. IEEE 802.11ac.
Dispositivo principal ejecutando la MS app (Set Top Box, STB)	PC1: Intel Celeron 1037U @ 1.80GHz x2, 4GB RAM, HDD 150GB, Ubuntu 14.04; Fast Ethernet embedded card; WiFi card IEEE 802.11 b/g/n; DVB-T USB card: Hauppauge! Nova-T Stick 3; TV LG 32LF592U (32").
Dispositivo secundario ejecutando la CS app	PC1, Samsung Galaxy Tab S (IEEE 802.11ac 10" Android 7 <i>tablet</i>), Samsung S5 (5" Android 7 <i>smartphone</i>)

7.3.1 PARTICIPANTES

Para la evaluación subjetiva de la plataforma con funcionalidades IDES básicas, han participado un total de 24 usuarios, de los cuales el 62% han sido hombres y solamente un 38% del total trabajan o estudian en áreas no-técnicas. La distribución por género, edad y el nivel de estudios se muestran en la Tabla 32.

Tabla 32 Perfil de los participantes

Género		Edad		Nivel de estudios	
Male	Female	<18	0%	Secundaria	12%
62%	38%	18-25	46%	F.P.	12%
		26-35	33%	Grado	63%
		36-45	4%	PhD	13%
		>45	17%		

7.3.2 METODOLOGÍA DE EVALUACIÓN

La evaluación subjetiva de la plataforma con funcionalidades IDES ha seguido las fases definidas al inicio de este capítulo. La fase 0 ha consistido en realizar a los participantes una breve introducción y contextualización del tema y de los objetivos de la investigación en curso. A continuación, durante las fases 1 y 2, los participantes han experimentado con el caso de uso implementado en la plataforma con el fin de realizar tests de sus funcionalidades para cambiar la reproducción entre las cuatro vistas disponibles y utilizando dos tipos de dispositivos secundarios disponibles (tablet Samsung Galaxy Tab S y smartphone Samsung Galaxy S5). Tras cada prueba (que se detallan a continuación) han rellenado un breve cuestionario, correspondiente con la fase 3 (ver Anexo 2).

El primer test ha consistido en analizar la utilidad e importancia de contar con un mecanismo de sincronización IDES entre un dispositivo principal y uno secundario. Para tal fin, los participantes han contado con una TV de 48" ejecutando la MS app y una Tablet de 10" ejecutando la CS app, ambas reproduciendo el mismo contenido¹⁸ (vista de la cámara frontal). Se han presentado 5 casos en los que se han forzado diferentes valores de asincronía entre los procesos de reproducción de los dos dispositivos involucrados, según se muestra en la Tabla 33.

¹⁸ Se ha decidido así ya que cuando ambos dispositivos presentan la misma vista, los usuarios perciben las asincronías más fácilmente que cuando utilizan diferentes vistas.

Tabla 33 Casos evaluados en el primer test

Caso	Descripción
1	Asincronía: -3s (CS app atrasada 3s respecto a la MS app).
2	Asincronía: -1s (CS app atrasada 1s)
3	Asincronía: 0s (sincronizado).
4	Asincronía: +1s (CS app adelantada 1s).
5	Asincronía: +3s (CS app adelantada 3s).

Cabe resaltar que el caso 3 se ha utilizado para evaluar el comportamiento adecuado de la plataforma, para el cual el participante debería obtener la mejor experiencia en cuanto a la sincronización entre dispositivos.

Los 5 casos han sido presentados a cada participante de forma aleatoria. En cada caso, se visualizó un vídeo durante 3 minutos, con el objetivo de que valoraran el nivel de sincronización alcanzado, y, en caso de percibir asincronías, si estas les resultaban molestas. Para ello, como se ha indicado, se utilizó la métrica MOS. Se han utilizado escalas de tipo Likert, para valorar el nivel de sincronización percibido, con 5 valores, del 1 al 5, siendo 1 la peor valoración posible y 5 la mejor valoración posible. En la Tabla 34 se exponen la escala utilizada para obtener el valor de la métrica aplicada tanto a la evaluación de la calidad del nivel de sincronización percibido como a la percepción y molestias ocasionadas por las (posibles) situaciones de asincronía detectadas durante el visionado de los vídeos, en el caso de que ocurrieran.

Tabla 34 Escala Likert utilizada para calcular la métrica MOS

Valor	Calidad de la sincronización	Efectos de las asincronías detectadas
5	Excelente	Imperceptible
4	Buena	Perceptible, pero no molesto
3	Neutral	Ligeramente molesto
2	Pobre	Molesto
1	Mala	Muy molesto

Tras finalizar el primer test, los usuarios han podido experimentar con la plataforma sin forzar ninguna asincronía entre los procesos de reproducción de la MS app y CS app, disponiendo de un dispositivo principal, ejecutando la MS app, en el que se presenta el contenido principal recibido por broadcast y de dos dispositivos (tablet y smartphone), ejecutando la CS app, en los cuales han podido cambiar la vista a visualizar (de entre las cuatro disponibles) obteniéndola a través de la red broadband. Tras acabar este segundo test, los participantes rellenaron un cuestionario, finalizando así la evaluación subjetiva.

Respecto a las condiciones de las redes de transmisión en este escenario, no se ha forzado pérdida de paquetes en la red broadcast ni retardos o limitaciones en el AB en la red broadband, puesto que el objetivo principal de la evaluación subjetiva ha sido, por un lado, valorar el nivel de sincronización percibido según las asincronías forzadas en cada uno de los cinco casos (Tabla 33) sin verse afectado por factores externos y, por otro lado, evaluar la funcionalidad y utilidad de la plataforma.

7.3.3 RESULTADOS OBTENIDOS

Respecto al primer test, en los que los participantes, de forma inconsciente, experimentan cinco niveles de sincronización diferentes (ver Tabla 34), los resultados obtenidos de la métrica MOS, junto con el I.C. 95%, se presentan en la Fig. 98.

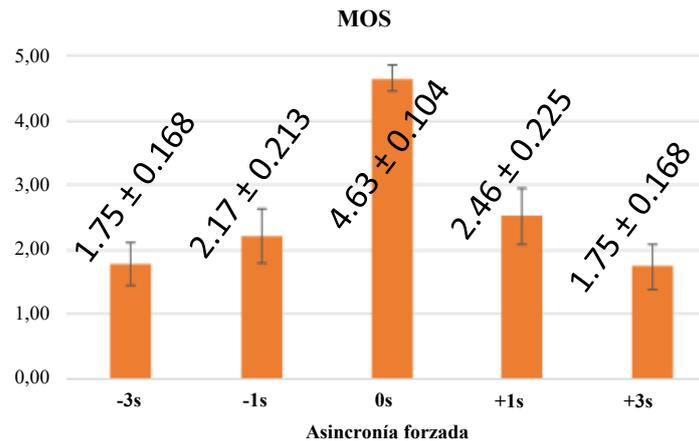


Fig. 98 Valor MOS de los cinco casos de asincronía forzada indicados en la Tabla 33

Se puede observar cómo la QoE percibida en el caso en el que no se fuerza ninguna asincronía (0s), la valoración MOS obtenida está próxima al máximo, que es 5. Además, se observa cómo la valoración MOS decae conforme aumenta la asincronía entre la reproducción de los dos dispositivos. Estos resultados confirman la importancia de proporcionar soluciones IDES precisas en escenarios multipantalla, como la diseñada en esta Tesis Doctoral

También puede observarse cómo los participantes consideran ligeramente más molesto el caso en el que el dispositivo secundario va atrasado respecto al dispositivo principal (caso -1s), comparado con el valor opuesto (caso 1s), en el que el nivel de sincronización alcanzado es el mismo, pero en este último caso el dispositivo secundario va adelantado.

HÁBITOS Y EXPERIENCIA PREVIA

Los participantes fueron preguntados acerca de sus hábitos de consumo, su experiencia previa en este tipo de escenarios y aplicaciones, el interés despertado y la utilidad y usabilidad de la plataforma. Los resultados obtenidos se muestran a continuación.

Por un lado, tal y como se puede ver en la Fig. 99, un 38% de los participantes indicó que nunca (21%) o casi nunca (16%) utiliza el móvil durante el consumo de TV para acceder a contenido o información relacionada con lo que está viendo. Por otro lado, los tipos de contenido más demandados para consumir en este tipo de escenarios fueron deportivos (19% de los participantes), vídeos musicales (17%), series (11%) y shows (11%).

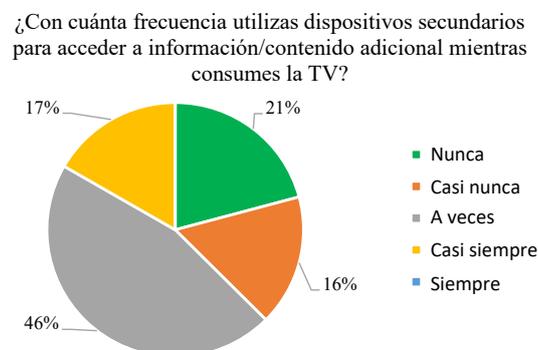


Fig. 99 Uso de dispositivos secundarios mientras se consume contenido de TV

Cuando los participantes fueron preguntados sobre si habían tenido alguna experiencia previa con escenarios o plataformas multipantalla, solamente el 25% afirmó haber experimentado algo similar, de los cuales el 67% aseguraron haber experimentado problemas relacionados directamente con la reproducción en segundas pantallas (debidos, por ejemplo, al retardo o la falta de sincronización entre los dispositivos), que les resultaron molestos.

SINCRONIZACIÓN INTER-DISPOSITIVO

Tras la fase en la que los participantes utilizan el escenario libremente, se han recogido los siguientes resultados acerca de la usabilidad y el comportamiento experimentado durante el uso del mismo. Para esta fase de la evaluación, también se ha utilizado otra escala del tipo Likert con los cinco niveles mostrados en las leyendas de las siguientes figuras, según el nivel de acuerdo (o desacuerdo) con cada una de las afirmaciones que se exponen en el cuestionario correspondiente. En las Fig. 100 hasta la Fig. 102, se muestran los resultados obtenidos. En la Fig. 100 se exponen los resultados acerca de la percepción del nivel de sincronización alcanzado entre los dispositivos principal y secundario. Un 58% de los participantes afirmó estar totalmente de acuerdo con afirmar que la sincronización alcanzada es precisa y un 38% estuvo parcialmente de acuerdo. Solamente un 4% mantuvo una opinión neutra respecto a esta afirmación. De hecho, los participantes también puntuaron sobre 10 la percepción del nivel de sincronización alcanzado, obteniendo como promedio un 9.

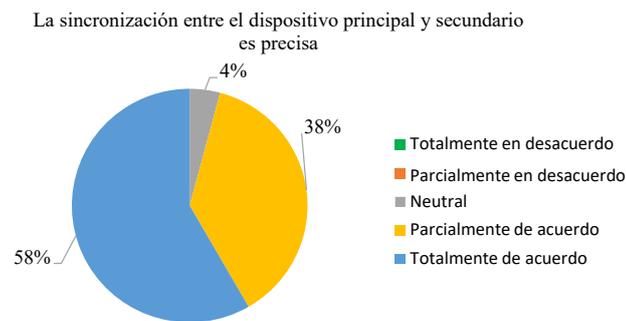


Fig. 100 Percepción del nivel de sincronización alcanzado entre los procesos de reproducción de los dispositivos principal y secundario

Por un lado, la Fig. 101 muestra la aceptabilidad del retardo existente al iniciar la reproducción en el dispositivo secundario (desde que se lanza por primera vez la CS app hasta que presenta contenido sincronizado). Solamente un 4% de los participantes no lo consideró aceptable. Por otro lado, la Fig. 102 muestra la aceptabilidad del retardo experimentado cuando, una vez la reproducción en el dispositivo secundario (en su CS app) ya está en ejecución, se cambia de cámara o vista, hasta que el nuevo contenido se presenta de forma sincronizada. En este caso, de nuevo, solamente un 4% de los participantes no consideró como aceptable dicho retardo. Aunque los retardos fueron percibidos por un número significativo de participantes, no lo consideraron como molestos, por lo que los resultados pueden considerarse bastante satisfactorios.

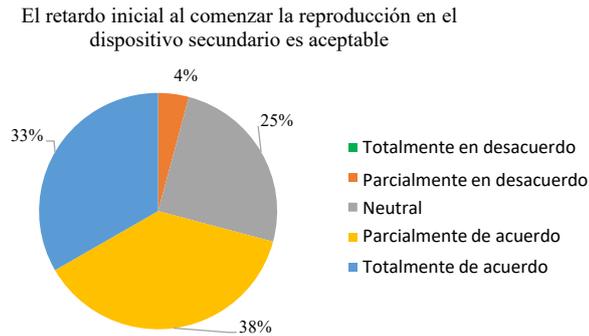


Fig. 101 Aceptabilidad del retardo inicial al lanzar la CS app en el dispositivo secundario

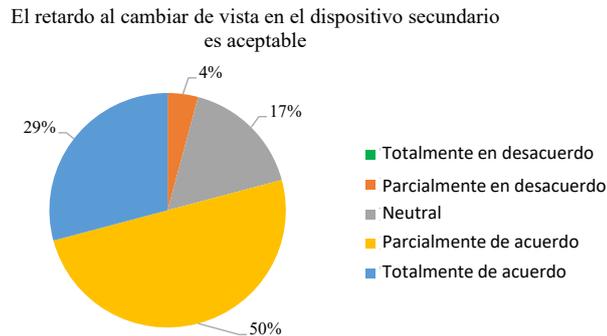


Fig. 102 Aceptabilidad del retardo al cambiar de cámara en la CS app en el dispositivo secundario

USABILIDAD, INTERÉS DESPERTADO Y APLICABILIDAD DE LA PLATAFORMA

A continuación, se exponen los resultados obtenidos acerca del interés despertado y de la aplicabilidad de la plataforma. A la pregunta de si esta plataforma puede proporcionar experiencias de consumo de TV más inmersivas, personalizadas y enriquecidas, el 83% afirmó estar totalmente de acuerdo; un 11% parcialmente de acuerdo; y el resto de los participantes se mostraron neutrales con esta pregunta.

La Fig. 103 muestra la opinión que tienen los participantes acerca del escenario evaluado, de los cuales un 96% lo consideró útil e interesante. La Fig. 104 muestra las opiniones acerca de la facilidad y comodidad de uso del escenario implementado en la plataforma desarrollada. Un 87% de los mismos la encontró cómoda y fácil de usar, y tan solo un 4% consideró lo contrario.

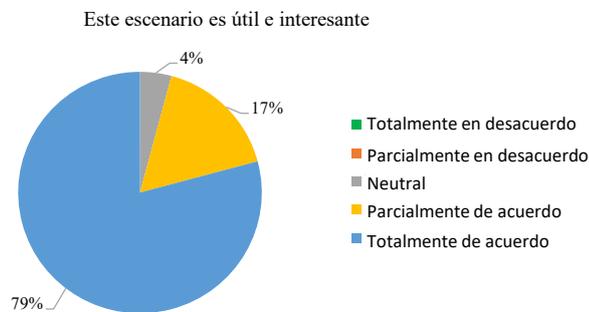


Fig. 103 Interés y utilidad del escenario evaluado

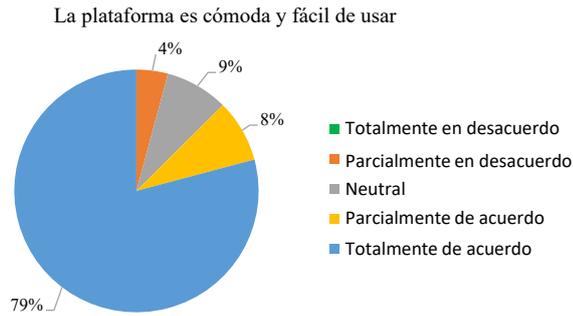


Fig. 104 Facilidad y comodidad de uso del escenario evaluado

Adicionalmente, la Fig. 105 muestra las opiniones de los participantes sobre si esta plataforma soluciona algún problema real. Un 83% de los participantes indicó que sí lo hace, frente a un 17% que tuvo una opinión neutral al respecto.

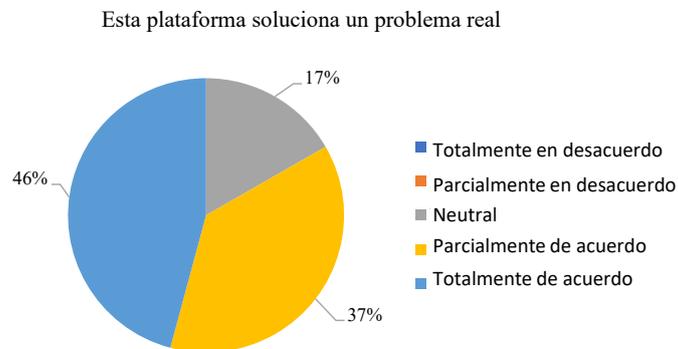


Fig. 105 Opinión acerca de si esta plataforma soluciona un problema real

La Fig. 106 muestra la opinión de los participantes sobre si esta plataforma permite mejorar la QoE experimentada durante el consumo de contenidos multipantalla. En este caso, la totalidad de los participantes así lo consideraron, siendo un 92% los que estuvieron totalmente de acuerdo y tan sólo un 8% los que estuvieron parcialmente de acuerdo.

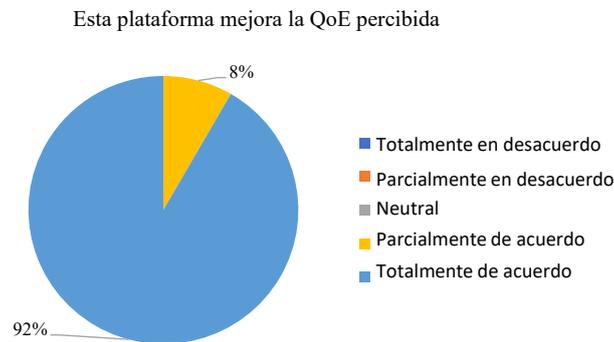


Fig. 106 Opinión sobre la mejora de la QoE percibida al utilizar la plataforma

Finalmente, la Fig. 107 muestra la opinión de los participantes acerca del posible impacto si esta plataforma estuviera disponible en el mercado. Un 96% considera que sí lo tendría y un 4%

se muestra neutral respecto a esta afirmación. Además, el 100% de los participantes utilizaría la plataforma si estuviera disponible.

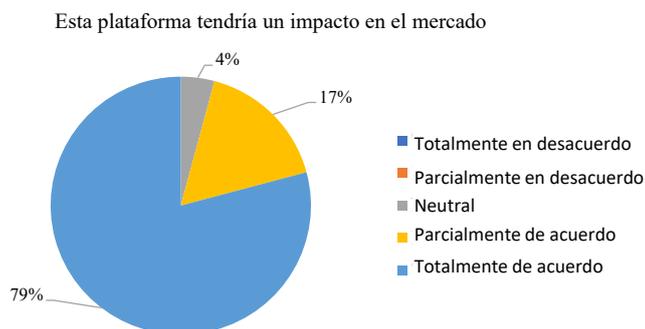


Fig. 107 Opinión acerca del posible impacto en el mercado de la plataforma

7.3.4 DISCUSIÓN

En este apartado se han expuesto los resultados obtenidos a partir de la evaluación subjetiva llevada a cabo sobre la plataforma con funcionalidades IDES, cuando se ha implementado en la misma un escenario con contenido multipantalla y multidispositivo. Los resultados obtenidos en la evaluación subjetiva son coherentes con los obtenidos en la evaluación objetiva mostrados en el capítulo anterior, ya que el nivel de sincronización alcanzado medido en esta última se corresponde con la buena valoración que han reflejado los participantes en la evaluación subjetiva. Además, los participantes en la evaluación subjetiva se mostraron satisfechos con la usabilidad y facilidad de uso de la plataforma, a pesar de que alguna de las respuestas evidencia que es necesario un desarrollo más profundo de las aplicaciones involucradas, con el fin de mejorar aspectos como, por ejemplo, los retardos de carga de contenido en los dispositivos secundarios (en las apps CS). Finalmente, se puede destacar el potencial prometedor de la plataforma a partir de las opiniones respecto a su adopción/adquisición si estuviera disponible en el mercado, o respecto a cómo ha afectado positivamente a la QoE percibida por los participantes.

7.4 ESCENARIO INTER-DISPOSITIVO INMERSIVO

En esta sección se va a evaluar la QoE percibida por los participantes en un escenario IDES como el anterior, pero al que se le ha añadido contenido omnidireccional. El contenido principal se reproduce en el dispositivo principal (ejecutando la MS app), la TV, mientras que el contenido secundario omnidireccional se puede reproducir en diferentes dispositivos secundarios (ejecutando la CS app) que soportan la reproducción de dicho tipo de contenido satisfactoriamente. En este caso, se han utilizado los tres dispositivos que mejores resultados han ofrecido en la evaluación objetiva (ver sección 6.3.3 del capítulo 6), como son el portátil, el smartphone Samsung Galaxy Note 9 (colocado en un HMD) y la tablet Samsung Galaxy Tab S4.

Respecto al contenido disponible, el contenido principal que se visualiza en el dispositivo principal, que es transmitido a través de la red broadcast, se ha extraído del contenido omnidireccional original. Además, se han generado diferentes contenidos con distinta temática (deportes, documentales y vídeos musicales), los cuales se listan en la Tabla 35, en la que se incluye la información necesaria para su correcta reproducción.

Tabla 35 Contenido generado para la evaluación subjetiva

Título	Duración (s)	Codificación	Vistas 2D	Vistas 360	Tipo de proyección vista 360	Fuente
F1 & Roller Coaster	360	H.264 + AAC	3	1	CMP	https://www.youtube.com/watch?v=fQoVFraBOnc&t=9s https://www.youtube.com/watch?v=8lsB-P8nGSM&t=89s
NBA Stephen Curry	143	H.264 + AAC	1	1	CMP	https://www.youtube.com/watch?v=y7tw6xz_h_0
Paris	271	H.264 + AAC	1	1	CMP	https://www.youtube.com/watch?v=EkshFcLESPU
Flight	684	H.264 + AAC	1	1	CMP	https://www.youtube.com/watch?v=HEElzZ7UjRg&t=159s
Sharks	198	H.264 + AAC	1	1	CMP	https://www.youtube.com/watch?v=j4ZT_XlhZBQ&t=5s
Shawn Mendes Live 1	303	H.264 + AAC	2	1	CMP	https://www.youtube.com/watch?v=LueM9tEu2wl https://www.youtube.com/watch?v=We7ZKgwqFKU
Shawn Mendes Live 2	240	H.264 + AAC	2	1	CMP	https://www.youtube.com/watch?v=ttByPyz3XFQ https://www.youtube.com/watch?v=rWx_m1VAp_E

7.4.1 PARTICIPANTES

El total de participantes en esta evaluación ha sido de 30 personas, de los cuales se proporciona información en la Tabla 36 (recogida durante la primera fase de la evaluación).

Tabla 36 Perfil de los participantes

Género		Edad		Nivel de estudios	
Male	Female	<21	0%	Secundaria	10%
57%	43%	21-25	37%	F.P.	10%
		26-30	14%	Grado	37%
		31-35	0%	Máster	13%
		36-40	13%	PhD	30%
		41-45	13%		
		>45	23%		

Cabe destacar que ninguno de los participantes presentaba deficiencias del tipo audiovisual, como, por ejemplo, problemas graves de visión o auditivos.

7.4.2 METODOLOGÍA DE EVALUACIÓN

Esta evaluación también se ha dividido en las fases definidas al inicio del capítulo. Previo al uso de la plataforma (fase 0), los participantes han recibido una breve introducción a la plataforma, y se les ha explicado sus principales características, así como los diferentes tipos de contenidos que iban a poder consumir.

Durante la fase 1, la evaluación se ha llevado a cabo en parejas (sean los usuarios A y B) y se han forzado las siguientes situaciones, con el fin de que todos los participantes hayan interactuado de todas las maneras posibles usando 3 dispositivos secundarios diferentes y durante periodos de 3 minutos:

- a) El usuario A consume el contenido en el dispositivo principal (MS app en TV) y el usuario B el contenido omnidireccional en el dispositivo secundario HMD (CS app);
- b) El usuario B consume el contenido en el dispositivo principal (MS app en TV) y el usuario A el contenido en el dispositivo secundario HMD (CS app);
- c) El usuario A consume el contenido complementario en la tablet como dispositivo secundario (en su CS app), el usuario B consume el contenido complementario en el portátil como dispositivo secundario (en su CS app), y ambos usuarios también pueden visualizar el dispositivo principal (con la MS app), situado enfrente de ellos y reproduciendo el contenido principal recibido por broadcast;
- d) El usuario B consume el contenido complementario en la tablet como dispositivo secundario (en su CS app), el usuario A consume el contenido complementario en el portátil como dispositivo secundario (en su CS app), y ambos usuarios también pueden visualizar el dispositivo principal (con la MS app), situado enfrente de ellos y reproduciendo el contenido principal recibido por broadcast.

A continuación, durante la fase 2, se ha permitido un uso libre del escenario: los participantes han podido escoger el dispositivo secundario que han preferido, mientras el dispositivo principal reproduce el contenido principal recibido por broadcast. Cuando han tenido que utilizar un dispositivo secundario (en la CS app) los participantes han tenido libertad para escoger qué contenido consumir (2D u omnidireccional) e interactuar con el escenario inmersivo utilizando cualquiera de los 3 dispositivos secundarios disponibles (portátil, tablet o HMD).

Tanto en la fase 1 como en la 2, los participantes han podido escoger el contenido a consumir de los listados en la Tabla 35.

Finalmente, durante la fase 3, los participantes han cumplimentado un cuestionario final para reflejar su opinión acerca de la experiencia global de consumo, la usabilidad del escenario implementado en la plataforma desarrollada, además de valorar el nivel de sincronización percibido entre el dispositivo principal (MS app) y los diferentes dispositivos secundarios (apps CS).

7.4.3 RESULTADOS OBTENIDOS

Tras utilizar el escenario (fases 2 y 3), los usuarios valoraron la experiencia de consumo en los diferentes dispositivos secundarios, tanto utilizando el salón virtual inmersivo, pero visualizando contenidos 2D en el TV virtual del mismo, como utilizando el visor de contenido omnidireccional. Los resultados de la métrica MOS \pm I.C. 95% se muestran en la Fig. 108.

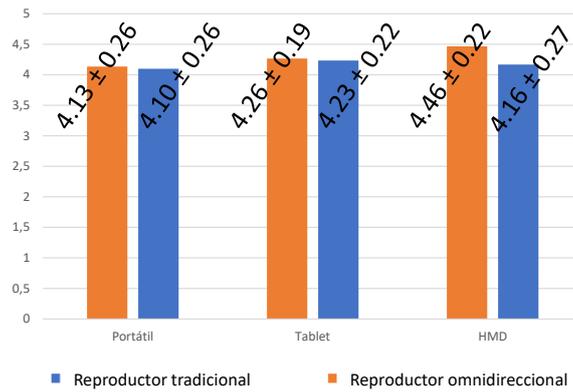


Fig. 108 Resultados MOS ± I.C.95% respecto a la experiencia de los participantes para cada dispositivo utilizado y respecto a la forma de consumir los contenidos relacionados

Los valores obtenidos son bastante similares para cada dispositivo secundario utilizado. Sin embargo, se puede observar cómo el uso del dispositivo HMD recibe una mejor valoración cuando se utiliza para consumir directamente contenido omnidireccional en el visor 360. Esto es comprensible, puesto que el HMD proporciona una experiencia de inmersividad completa, siendo el usuario capaz de ver el contenido 360 por completo sin posibilidad de ver nada más del entorno real a su alrededor. Mediante el uso de dicho dispositivo se permite al usuario navegar por los 360 grados del contenido multimedia de manera natural (con el movimiento de la cabeza). Además, un 80% de los participantes aseguraron no haber notado ningún problema durante la experiencia. Del 20% restante, los problemas experimentados fueron respecto a la calidad del contenido al cambiar de vista, lo cual es explicable debido a la política del reproductor de contenido DASH implementado en la plataforma, que selecciona la calidad más baja para los segmentos iniciales del contenido.

HÁBITOS Y EXPERIENCIA PREVIA

Respecto a la información recogida acerca de los hábitos y experiencia previa de los participantes en esta evaluación con algún sistema o aplicación similar, a continuación, se resumen los resultados obtenidos. La Fig. 109 y la Fig. 110 muestran, respectivamente, la cantidad de horas que dedican los participantes a consumir contenido multimedia online (por ejemplo, a través de aplicaciones como Netflix, HBO, YouTube, etc.) o broadcast (por ejemplo, programas de Antena 3, TVE, etc. en la TV).

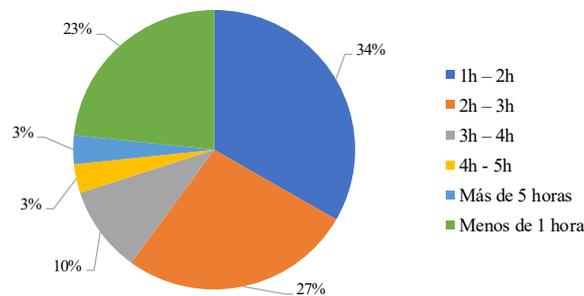


Fig. 109 Horas diarias utilizadas para consumir contenido online

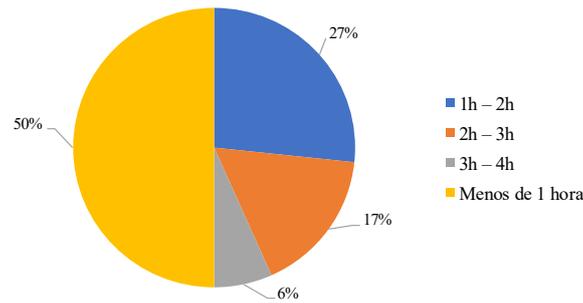


Fig. 110 Horas diarias utilizadas para consumir contenido broadcast de TV

Un 23.3% de los participantes afirmó consumir menos de 1 hora diariamente de contenido online, mientras que, cuando se trata de contenido de TV tradicional (broadcast), el porcentaje de participantes lo consumieron durante menos de 1 hora diaria sube hasta el 50%. La Fig. 111 muestra el porcentaje de participantes que usa cada uno de los dispositivos listados para el consumo de contenido online, a través de la red broadband.

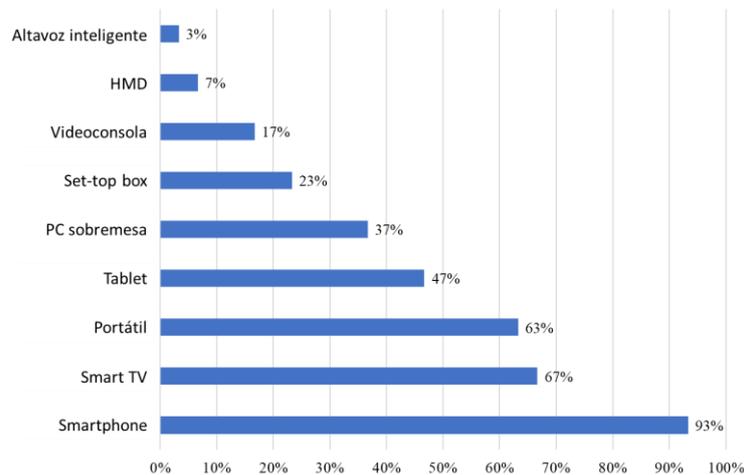


Fig. 111 Uso por parte de los participantes de diferentes dispositivos para el consumo de contenido online

La gran mayoría de los participantes (93%), afirmaron utilizar el smartphone para acceder a contenido multimedia online. Sin embargo, solamente un 7% de los participantes consumen contenido a través de dispositivos HMD. En la Fig. 112 se muestra la experiencia previa de los participantes en cuanto al uso de dispositivos HMD.

Más de la mitad de los participantes (57%) no han tenido ninguna experiencia previa con este tipo de dispositivos y solamente el 10% posee, al menos, un HMD. En concreto, los que han tenido alguna experiencia previa con dispositivos HMD, ha consistido en¹⁹: consumir contenido audiovisual (61.5%), ejecutar aplicaciones de RV (46.2%) y en videojuegos de RV (30.8%). Cuando los participantes fueron preguntados sobre si creen que este tipo de dispositivos puede enriquecer la experiencia de consumo de contenidos multimedia, el 84.6% afirmó estar de acuerdo.

¹⁹ La pregunta realizada es del tipo de respuesta múltiple (multiopción)



Fig. 112 Experiencia previa de los participantes con el uso de dispositivos HMD

SINCRONIZACIÓN INTER-DISPOSITIVO

Durante la fase 2 de uso guiado del escenario, cuando los participantes utilizaron el portátil o la tablet como dispositivo secundario, pudiendo visualizar simultáneamente el contenido principal presentado en el TV (MS app). En dichos casos, un 86.7% (cuando estaban utilizando el portátil) y un 90% (cuando estaban utilizando la tablet) de los participantes aseguraron percibir la reproducción de los contenidos en el dispositivo principal y el secundario (en MS y CS apps) perfectamente sincronizada. Además, respecto a la visualización del contenido complementario a través del portátil, el 60% de los participantes prefirieron la visualización de contenido omnidireccional respecto al tradicional. Respecto a la visualización del contenido complementario a través de la tablet, dicha preferencia se incrementó hasta un 80%. En ambos casos (utilizando como dispositivo secundario el portátil o la tablet), los participantes prefirieron visualizar antes el contenido del dispositivo secundario que el del principal (70% y 80%, respectivamente).

Cuando uno de los participantes de la pareja consumió el contenido a través de la TV (MS app) y el otro, a través del HMD como dispositivo secundario (CS app), se obtuvieron las valoraciones que se exponen a continuación. Respecto a los usuarios utilizando el HMD, un 90% afirmó no haberse sentido aislado y no consideró que dicho dispositivo le impidiera comunicarse con el otro participante durante la experiencia de consumo compartida. Respecto a los usuarios visualizando la TV, un 76.6% de ellos afirmó no haberse sentido aislados y que pudieron comunicarse con el participante que estaba utilizando el HMD. Respecto a la sensación de estar ambos sincronizados, un 96.7% de los participantes que utilizó el HMD afirmaron tener la sensación de haber consumido el contenido de forma perfectamente sincronizada con el contenido que estaba visualizando el otro participante, y un 83.3% de los participantes que visualizó la TV respondieron de la misma manera. Sin embargo, un 16.6% de los usuarios que utilizaron el HMD se sintieron incómodos durante la experiencia de consumo, ya que se tuvieron que mover físicamente para ver las distintas regiones de las escenas 360 del contenido. De hecho, un 20% de dichos usuarios afirmaron haber necesitado ponerse de pie en algún momento. De manera análoga, un 50% de los participantes que visualizó la TV se dieron cuenta de que su compañero se estuvo moviendo, pero no les resultó incómodo ni molesto. Solamente a un 3.3% les resultó algo molesto. Además, respecto a los participantes utilizando el HMD, un 53.3% afirmó haber sentido que el tiempo les pasó más rápido.

Respecto a la fase 3, correspondiente al uso libre del escenario, un 73.7% de los participantes afirmaron no haber notado diferencias entre los instantes de reproducción del dispositivo principal (MS app) y los de los diferentes dispositivos secundarios (apps CS). Adicionalmente, un 94.7% consideró que el hecho de cambiar de dispositivo secundario durante la sesión no le afectó a la experiencia de consumo.

USABILIDAD, INTERÉS DESPERTADO Y APLICABILIDAD DEL ESCENARIO

Respecto a la preferencia en la elección del dispositivo de consumo durante la sesión, el 53% de los participantes prefirió el HMD, un 53% la tablet y un 16% el portátil. Además, el 76.7% aseguró que, durante la fase de uso libre, no quisieron ver ni vieron el contenido en el dispositivo principal (TV con MS app). El resto (23.3%) que sí lo visualizó fue para comprobar que el contenido que se presentaba en la CS app estaba sincronizado con el contenido de la MS app y que el nivel de sincronización alcanzado era satisfactorio.

Respecto a los participantes que prefirieron el HMD durante algún momento de la evaluación, aseguraron que este les ha proporcionado una experiencia más realista e inmersiva. Los participantes que seleccionaron la Tablet en algún momento afirmaron que este dispositivo les proporcionaba una forma de interactuar más fácil y que les resultó más cómodo para la vista. Finalmente, los participantes que seleccionaron el portátil en algún momento declararon que este dispositivo les permitió tener un mayor control sobre el escenario.

Respecto a la usabilidad global de las aplicaciones del caso de uso implementado en este escenario, los participantes respondieron al cuestionario SUS [169], que consta de las 10 preguntas que aparecen en la Tabla 37.

Tabla 37 Cuestionario SUS

Q1	Creo que me gustaría utilizar este sistema frecuentemente
Q2	He encontrado el sistema innecesariamente complejo
Q3	El sistema era fácil de usar
Q4	Creo que necesitaría ayuda técnica para utilizar este sistema
Q5	Las funciones de este sistema están bien integradas
Q6	Creo que había demasiada inconsistencia en el sistema
Q7	Creo que la mayoría de las personas aprenderían a utilizar este sistema de forma rápida
Q8	El sistema ha sido engorroso de utilizar
Q9	Me sentí con seguridad utilizando el sistema desde el principio
Q10	Necesité aprender muchos conceptos antes de usar el sistema

Los resultados de las respuestas a dichas preguntas se pueden observar en la Fig. 113, y proporcionan una puntuación agregada total del 0 al 100, siendo 0 el peor resultado posible y 100 el mejor, tal y como está definido en [170]. En este caso, la puntuación obtenida fue de 92.67 puntos, con una desviación estándar de ± 13.17 . De acuerdo con [170], se puede afirmar que la usabilidad que proporciona el escenario evaluado en la plataforma desarrollada en la Tesis Doctoral, se puede considerar como 'excelente'. No obstante, a los participantes también se les preguntó por cualquier cambio o sugerencia con el fin de mejorar la experiencia del usuario, y las observaciones recopiladas están relacionadas principalmente con el dispositivo de consumo, especialmente cuando se utiliza el HMD. En particular, el 6% de los participantes afirmaron haberse sentido mareados en algún momento puntual de la experiencia de consumo y prefirieron interactuar con otros dispositivos secundarios, como la tablet o el portátil.

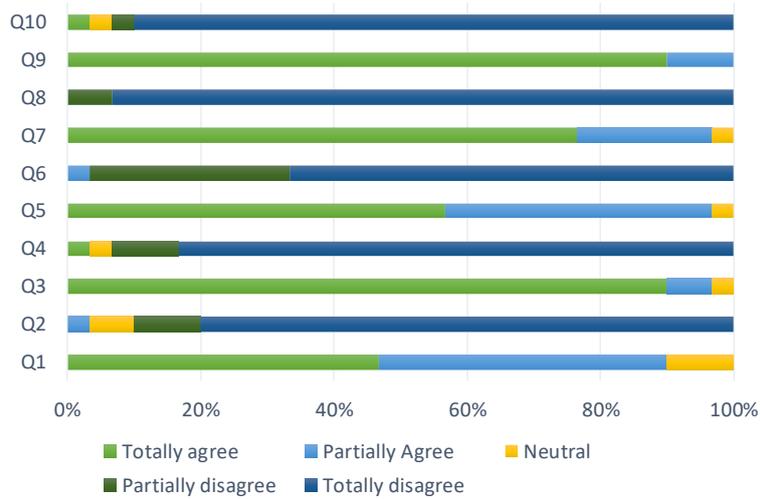


Fig. 113 Resultados del cuestionario SUS

Respecto a la interfaz de usuario del visor web para el consumo de contenido omnidireccional (salón inmersivo), los participantes sugirieron cambios menores (por ejemplo, mayor tamaño de la TV virtual) u otros cambios más significativos como cambiar la ubicación de los botones con los que se interacciona a través del dispositivo HMD, para hacerlos menos intrusivos (originalmente ubicados en la zona central inferior del FoV, lo que podría ocultar contenido relevante).

A pesar de las observaciones de los usuarios, se puede concluir que la usabilidad proporcionada a los participantes en este escenario es muy satisfactoria. Según los resultados mostrados en la Fig. 114 y en la Fig. 115, un 93% de los participantes afirmó sentirse más integrado/inmerso en la experiencia de consumo de contenido omnidireccional, e incluso un 67% de los participantes afirmó haber perdido la noción del tiempo durante la experiencia.



Fig. 114 Inmersividad percibida por los participantes durante la evaluación

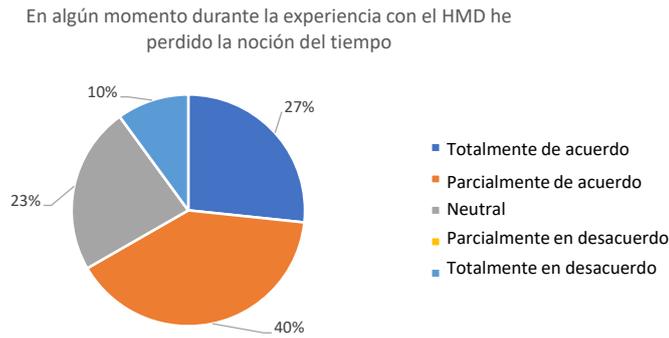


Fig. 115 Porcentaje de participantes que han perdido la noción del tiempo en algún momento

Del total de participantes, un 93.4% opinó que incluir contenido omnidireccional en servicios de TV híbrida proporciona una experiencia divertida y al 86.7% no le importaría repetir la experiencia. Por otro lado, la Fig. 116 muestra cómo de acuerdo estuvieron los participantes con la siguiente afirmación: *“Considero que la inclusión de contenidos omnidireccionales (360) en los servicios de TV contribuirá a que los usuarios se sientan más enganchados a los canales que ofrezcan dicho tipo de servicios (es decir, que no hagan zapping rápidamente y que se queden más tiempo visualizando el canal correspondiente)”*.

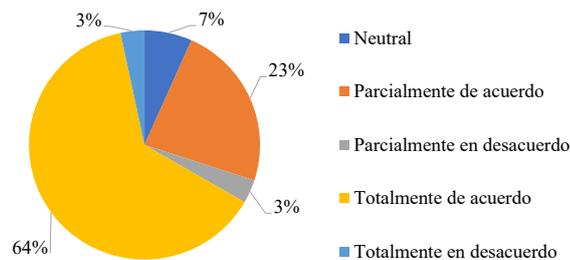


Fig. 116 Opinión respecto a la mejora que proporciona a la experiencia de consumo de TV la inclusión de contenidos omnidireccionales complementarios

De nuevo, pero tras haber experimentado con este escenario, los participantes dieron su opinión respecto a la afirmación *“Creo que este tipo de contenidos mejora la experiencia del usuario de TV”*. El 84.6% estuvo de acuerdo con dicha afirmación antes de experimentar con el mismo. Tras su uso durante la evaluación, el porcentaje que pasó a estar de acuerdo con dicha afirmación aumentó hasta el 96.7%. Además, un 96.6% consideró que el tiempo de aprendizaje para utilizar el escenario es corto, y un 93.3% consideró que está suficientemente familiarizado con los dispositivos utilizados durante la evaluación y no tendría problemas en integrarlos en su hogar si la posibilidad de crear un escenario así estuviera disponible comercialmente. Para finalizar, la Fig. 117 muestra para qué tipo de contenidos utilizarían este escenario de uso implementado en la plataforma. Cabe resaltar que la gran mayoría de participantes ha seleccionado los géneros de documentales (90%) y eventos deportivos (87%).

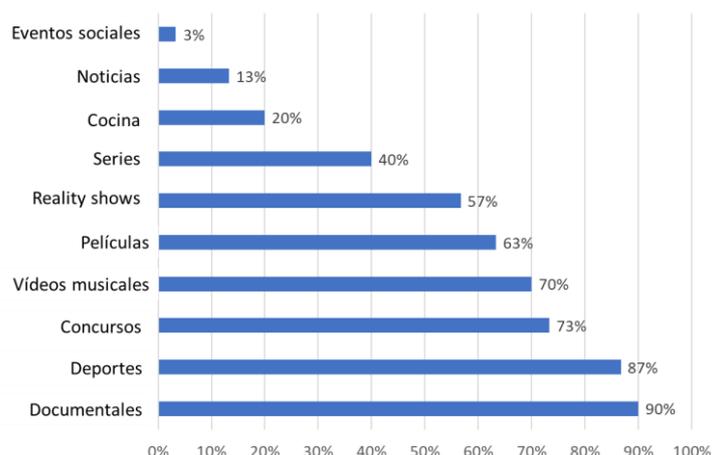


Fig. 117 Temática o género del contenido para cuyo consumo los participantes utilizarían este escenario implementado en la plataforma desarrollada en la Tesis Doctoral

7.4.4 DISCUSIÓN

En la evaluación de este escenario en el que, además, se han integrado todas las funcionalidades necesarias para poder consumir contenido omnidireccional, se ha comprobado el buen funcionamiento de los componentes de la plataforma desarrollada en la Tesis Doctoral. En vista a los resultados obtenidos, el nivel de sincronización alcanzado ha sido satisfactorio y no ha afectado negativamente a la experiencia de los participantes. Cabe destacar que para la evaluación se ha tenido que optar por utilizar únicamente los dispositivos secundarios que mejor rendimiento habían ofrecido en la evaluación objetiva, ya que los contenidos omnidireccionales necesitan ser reproducidos por dispositivos modernos que dispongan de buenas prestaciones. Tras comprobar el interés despertado y la buena valoración respecto a la usabilidad proporcionada por el escenario y las aplicaciones implementadas, se puede concluir que el escenario implementado en la plataforma desarrollada en la Tesis Doctoral con funcionalidades IDES y contenido omnidireccional (inmersivo) es capaz de proporcionar resultados muy satisfactorios y ofrecer una buena QoE a los participantes.

7.5 ESCENARIO INTER-DESTINATARIO

Con el fin de evaluar la percepción de los usuarios del funcionamiento de los mecanismos de sincronización IDMS integrados en la plataforma, se ha implementado en la misma un escenario de TV Social, incluyendo una herramienta de interacción social durante la sesión compartida, mediante mensajes de texto.

Para la evaluación subjetiva, se han habilitado cuatro puntos de consumo distribuidos en el campus universitario, cada uno emulando un hogar e incluyendo dos dispositivos, uno principal (terminal híbrido) y otro secundario. Los dispositivos empleados para cada uno de los cuatro puntos de consumo son los mismos que los utilizados en la evaluación objetiva (ver sección de evaluación IDMS del Capítulo 6). La Tabla 38 resume los dispositivos empleados.

Tabla 38 Dispositivos involucrados en la evaluación del escenario IDMS

Dispositivo	Características
PC _B (<i>broadcaster</i>)	Intel Core i7-6700 @ 3.40GHz, 8GB RAM, SSD 240GB, Windows 10
Transmisor DVB-T	DVB-T PCI Card: DekTec DTA-2111 + DekTec StreamXpress SW
Distribución DVB	Tecatel SMATV Training Panel (con amplificadores, divisores de señal, etc.).
PC _S (servidor multimedia, NTP y Manager IDMS)	Intel Xeon E5420 @ 2.50GHz x8, 8GB RAM, Ubuntu 14.04, Apache Server v2.4.7, Fast Ethernet embedded card.
Routers	4xR: TP-Link AC1900 Wireless Dual Band Gigabit. IEEE 802.11ac.
Hybrid Terminal (Set Top Box, STB)	4xPC _H : Intel Celeron 1037U @ 1.80GHz x2, 4GB RAM, Ubuntu 14.04; Fast Ethernet embedded card; WiFi card IEEE 802.11 b/g/n
Receptor DVB	DVB-T USB card.with RTL2832u chipset
TVs x4 (monitores)	TV1: LG 32LF592U (32"); TV2: Panasonic TH42PX80E (42"); TV3 & TV4: Samsung Smart TV UE48H6200AWXXC (48").
Dispositivos secundarios (apps CS)	2x Samsung Galaxy Tab S (IEEE 802.11ac 10" tablet) 1x Samsung S5 (IEEE 802.11ac 5" smartphone) 1x Samsung S6 (IEEE 802.11ac 5" smartphone)

Para este caso particular de TV Social, el contenido multimedia emitido a través de DVB-T ha sido al que le ha sido aplicado el algoritmo de ajuste de reproducción IDMS, puesto que esta es la tecnología tradicional para la transmisión de contenido en entornos de TV en los hogares. Por lo tanto, durante la evaluación los participantes consumirán el contenido reproducido en el dispositivo principal (terminal híbrido), a través de la MS app, de forma sincronizada.

Por un lado, en este caso se ha adoptado un esquema de control de la sincronización IDMS centralizado o SMS, puesto que, tal y como se ha visto en el capítulo anterior, presenta mejores resultados respecto al nivel de sincronización IDMS alcanzado que con el esquema de control distribuido DCS. Por otro lado, como referencia de sincronización se ha seleccionado el punto de reproducción de la MS app más atrasada. De esta manera, se pretende valorar hasta qué punto seleccionar esta estrategia (que requiere de un número de ajustes mayor que si la referencia es un punto de reproducción promedio) perjudica o afecta a la QoE percibida por los usuarios involucrados en la evaluación. Durante esta evaluación, tampoco se han provocado pérdidas de paquetes en la transmisión broadcast ni retardos en la red broadband, por el mismo motivo que en la anterior.

Respecto al contenido utilizado, se ha preparado un contenido de temática deportiva, concretamente un partido de baloncesto de la NBA. En esta evaluación se ha incluido únicamente contenido multimedia (codificado como H.264 + AAC, con una tasa de 25fps y encapsulado en flujos MPEG2-TS) para ser transmitido vía broadcast. Además, a través de la red broadband, se ha establecido una herramienta de chat sincronizada (CS app) cuyo diseño e implementación se expone a continuación.

En concreto, respecto a la evaluación de este escenario, se ha buscado determinar los siguientes puntos:

- La satisfacción de los participantes respecto al rendimiento y usabilidad de la plataforma;
- La opinión de los participantes respecto a la aplicabilidad de la plataforma y la utilidad de sus funcionalidades;
- El interés despertado en los participantes acerca de este tipo de plataformas o escenarios de TV Social;

- Si los algoritmos IDMS y los escenarios de TV Social proporcionan al participante un mayor sentimiento de cercanía para con el resto de los participantes en la sesión de consumo, si les hace estar más atentos al contenido y si se sienten más inmersos en el consumo del contenido.

7.5.1 INTERFAZ DE MENSAJES CONTEXTUALIZADOS POR EVENTO EN EL CANAL DE CHAT

Respecto a los mensajes contextualizados disponibles para el servicio de chat, al tratarse de un partido de baloncesto, se han añadido para cada uno de los equipos, mensajes referentes a: meter un triple, canasta de 2 puntos, fallo de equipo, canasta de tiro libre, pérdida de balón, falta, robo de balón y tiempo agotado.

En la Fig. 118 se muestra la diferencia del aspecto de la pantalla del terminal híbrido cuando está visible (o no) la herramienta de chat. La Fig. 119 muestra el aspecto de la CS app de la herramienta de chat ejecutada en un dispositivo secundario.

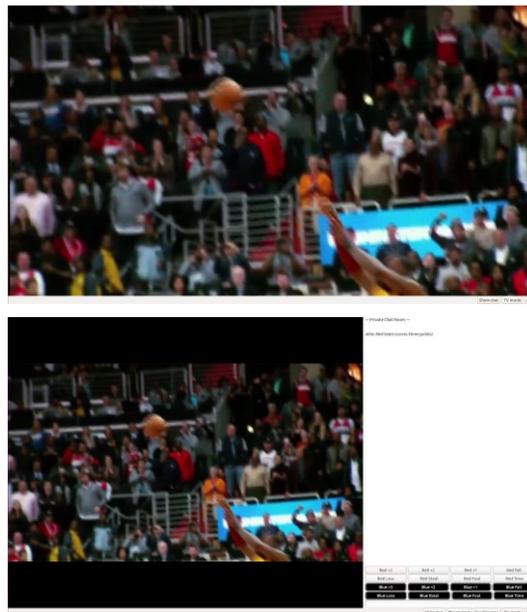


Fig. 118 Aspecto de la pantalla del terminal híbrido con el contenido visualizado en la MS app a pantalla completa (superior) o junto con la CS app de la herramienta de chat (inferior)

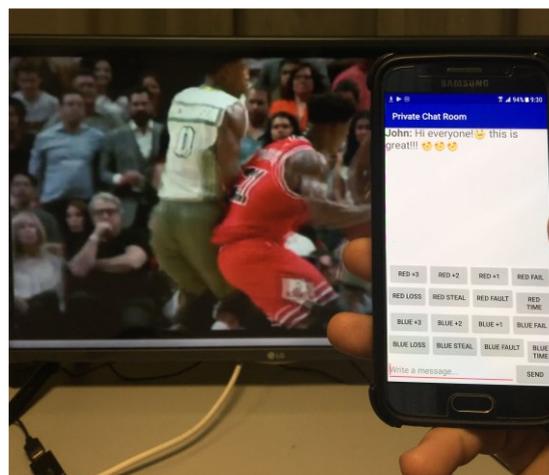


Fig. 119 CS app de la herramienta de chat en dispositivo secundario (smartphone)

7.5.2 PARTICIPANTES

Para esta evaluación, se ha contado con un total de 59 participantes, aunque 6 de ellos han sido descartados por las incoherencias encontradas en sus respuestas. La Tabla 39 muestra el perfil de los mismos, su género, rango de edad y nivel de estudios. Ninguno de los participantes presentó ningún tipo de problema visual o auditivo que hubiera podido afectar al proceso de evaluación.

Tabla 39 Perfil de los participantes

Género		Edad		Nivel de estudios	%
Mujer (%)	Hombre (%)	Años	%		
51	49	<18	5.66	Estudios obligatorios / instituto	5.66
		18-24	33.96	Estudiante universitario	45.28
		25-29	30.19	Graduado / Máster	30.19
		30-39	9.43	PhD	18.87
		40-49	18.87		
		50-59	1.89		

7.5.3 METODOLOGÍA DE EVALUACIÓN

Tras la fase 0, correspondiente con la cumplimentación del cuestionario previo para recoger la información de los participantes, y la fase 1, correspondiente con el proceso de aprendizaje y familiarización con la plataforma y el escenario en cuestión, se lleva a cabo la fase 2, la cual se divide, en este caso, en dos etapas.

La primera etapa de la fase 2 ha consistido en la evaluación, por pares, de cinco casos en los que se ha forzado (fijado) un valor de asincronía concreto entre los procesos de reproducción de las apps MS de los dispositivos principales de cada hogar. Esto permite analizar el impacto de la falta de sincronización en experiencias de consumo compartidas (en este caso, en escenarios del caso de uso de TV Social).

En concreto, los valores de asincronía fijados entre los procesos de reproducción de las MS app de las parejas de participantes han sido de -4s, -2s, 0s (sincronizado), 2s y 4s. Para evidenciar esta falta de sincronización entre los participantes, los mecanismos de sincronización del chat (ver sección 5.2.3) también se han deshabilitado. Posteriormente, los participantes han valorado la QoE percibida empleando una escala de tipo Likert, obteniendo valores de la métrica MOS [168].

Para la segunda etapa de la fase 2 de la evaluación, se han habilitado sesiones de consumo de TV compartidas entre cuatro participantes, los cuales han podido utilizar el escenario libremente. La única condición preliminar para esta fase de la evaluación ha sido que los participantes que forman parte de cada sesión se conocieran previamente, con el fin de facilitar una comunicación fluida y natural durante la experiencia de consumo multimedia. Los principales objetivos de esta parte eran determinar: 1) si los participantes habían tenido alguna experiencia previa similar a la ofrecida por el escenario implementado en este caso en la plataforma; 2) la usabilidad y aplicabilidad de dicho escenario; y 3) si el valor del umbral del nivel de sincronización IDMS (*session async*) configurado era suficiente para proporcionar una experiencia satisfactoria a los

participantes. Respecto a las condiciones de transmisión del contenido broadcast, se han configurado una serie de retardos iniciales para cada uno de los MS involucrados de 0s, 1.5s, 3s y 6s. De esta manera se emulan diferencias de retardos en la recepción del contenido broadcast entre los 4 hogares y se fuerza a realizar una sincronización inicial entre las MS apps de cada hogar.

7.5.4 RESULTADOS OBTENIDOS

En la segunda fase de la evaluación, que corresponde con un escenario en el que los participantes evaluaron el escenario en parejas (cada uno en un "hogar") y evaluaron, utilizando una escala de tipo Likert de 5 niveles, la percepción de sincronización percibida bajo diferentes condiciones de sincronización desconocida para ellos. Una vez los participantes estaban seguros de qué puntuación otorgar al caso experimentado, lo notificaron al supervisor y rellenaron un cuestionario. Este procedimiento se repitió para los cinco casos propuestos (con valores de asincronía forzados entre las MS apps de cada hogar de -4s, -2s, 0s, 2s y 4s). Los resultados de MOS obtenidos se muestran en la Fig. 120, incluyendo el valor medio y el I.C. 95% para cada caso. Los valores negativos y positivos en el eje X significan que el proceso de reproducción de la MS app del participante va adelantado o atrasado, con respecto al proceso de reproducción de la MS app de su compañero.

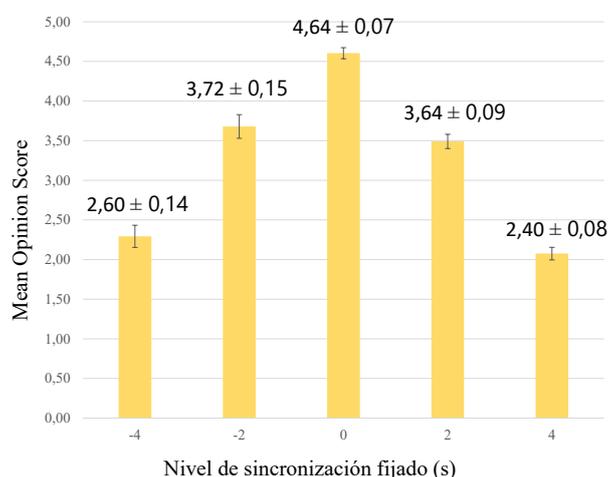


Fig. 120 Valor MOS e I.C. 95% obtenido en la segunda fase de la evaluación subjetiva

Tal y como se puede observar, un aumento en el valor de asincronía afectó negativamente a la QoE, siendo percibida a partir de los $\pm 2s$ y llegando a ser molesta cuando se superó los $\pm 4s$. Como era de esperar, la mejor valoración se obtiene cuando no existió asincronía (0s), es decir, cuando los dos participantes de la pareja consumieron el contenido de manera perfectamente sincronizada. Además, se puede observar que cuando un participante consumió el contenido atrasado respecto al otro participante de la pareja (asincronía > 0s), evaluó con una puntuación peor que cuando estaba adelantado (asincronía < 0s). Esto puede deberse a la posibilidad de haber recibido algún mensaje de chat que anticipara algún evento del contenido (*spoiler*). Como conclusión, se puede deducir que, cuanto más alto es el valor de asincronía entre los hogares o participantes involucrados, peor es la calidad de la experiencia (QoE) percibida por el mismo y que, además, para el caso en el que el proceso de reproducción del participante u hogar esté por detrás del resto de participantes, la QoE percibida será peor.

Respecto a la tercera fase de la evaluación, se ha podido obtener una gran cantidad de información relevante a partir de las respuestas proporcionadas por los usuarios. El propósito de esta fase ha sido recoger la mayor cantidad posible de información acerca de experiencias previas similares que hayan tenido los participantes, así como los problemas que hayan podido

experimentar durante las mismas. Asimismo, como se puede ver a partir de los resultados obtenidos, se confirma, por un lado, que este tipo de escenarios implementados en la plataforma desarrollada en la Tesis Doctoral, despiertan el interés de los usuarios y, por otro lado, la aceptación que plataformas como la desarrollada en la Tesis Doctoral para ofrecer este tipo de servicios podrían tener en el mercado. En esta fase de la evaluación, se ha segmentado por género la información obtenida, con el objetivo de conocer si existe alguna diferencia entre géneros y para analizar el impacto que ofrece este tipo de escenarios de manera más detallada. Los resultados obtenidos se han dividido en cuatro partes: 1) hábitos y experiencia previa; 2) percepción de la sincronización IDMS; 3) usabilidad y comportamiento del escenario implementado en la plataforma; y 4) interés despertado y aplicabilidad de dicho escenario.

HÁBITOS Y EXPERIENCIA PREVIA

Durante la tercera fase de la evaluación, los participantes fueron preguntados acerca de si habían tenido alguna experiencia previa con este tipo de escenarios y acerca de sus hábitos de consumo. Solamente un 25% de los participantes afirmó haber utilizado en algún momento una aplicación o escenario similar al presentado (Fig. 121), de los cuales un 62% afirmó haber sufrido situaciones molestas (Fig. 122), relacionadas, principalmente, con *spoilers* en el canal de interacción social o demasiados participantes en dicho canal, al tratarse de chats con acceso abierto y no privados (a diferencia de la herramienta de chat implementada en la plataforma desarrollada para poder interactuar socialmente, de manera privada, en escenarios que proporcionen experiencias compartidas entre usuarios remotos, como es el caso de escenario de TV social evaluado en este apartado).

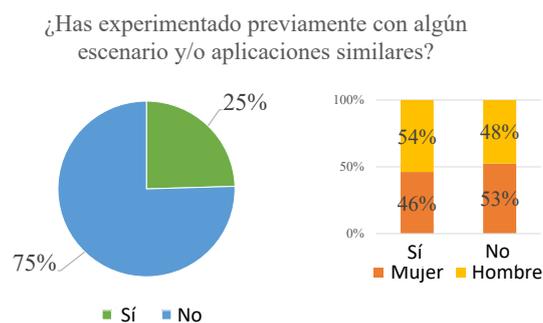


Fig. 121 Experiencia previa con escenarios o aplicaciones similares

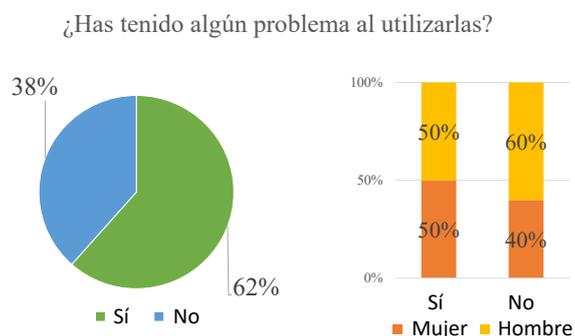


Fig. 122 Problemas con plataformas similares

Más del 80% de los participantes suelen quedar, en algún momento, con amigos o familiares en algún lugar (hogar, bar...) para consumir eventos o programas específicos en la TV (Fig. 123). Solamente el 13% de los participantes afirman que nunca han interactuado con otras personas (remotas) durante el consumo de TV (Fig. 124). Además, un 77% de los participantes opinó que

la distancia geográfica entre amigos o familiares es un impedimento para consumir y compartir este tipo de experiencias (Fig. 125).

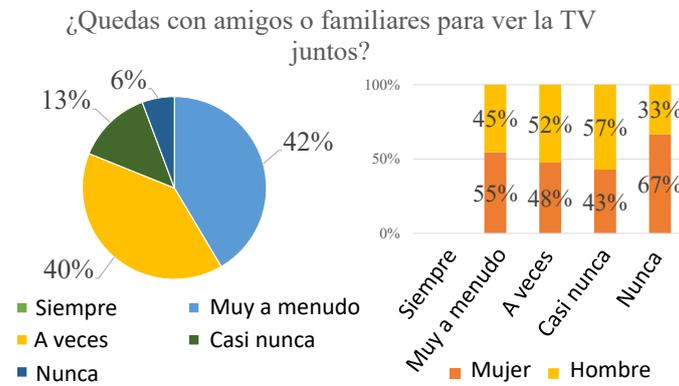


Fig. 123 Respuestas acerca del consumo compartido de contenido de TV en un mismo lugar



Fig. 124 Respuestas acerca del consumo compartido de contenido TV en diferentes lugares (TV Social)

¿La distancia te ha impedido disfrutar de este tipo de experiencias compartidas?

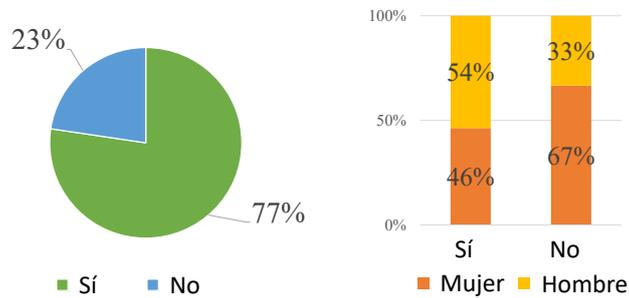


Fig. 125 Percepción de la distancia geográfica como impedimento para tener este tipo de experiencias compartidas

Respecto a si consideran que el uso de este tipo de aplicaciones o escenarios (TV social) mejora la QoE percibida durante el consumo de contenidos de TV, el 88% respondió de manera afirmativa (Fig. 126).

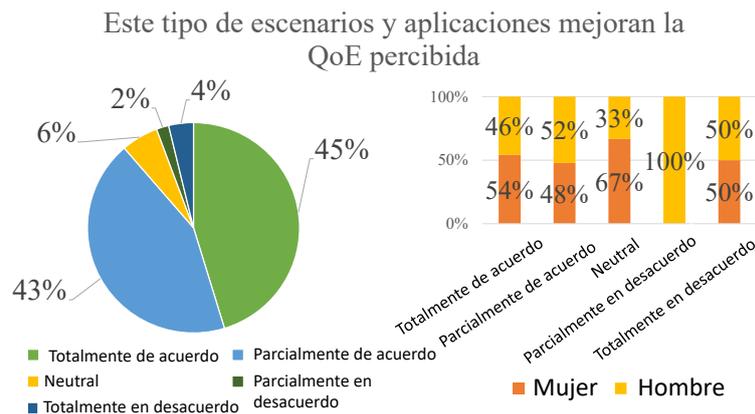


Fig. 126 Opinión respecto a cómo afecta a la QoE percibida por los usuarios este tipo de aplicaciones o escenarios

Desde el punto de vista de la clasificación de los resultados por género, las respuestas de hombres y mujeres fueron similares, no apreciándose diferencias significativas. Independientemente del género, los participantes tienen hábitos similares en cuanto al consumo de contenido de TV y experimentaron problemas similares cuando han querido compartir el consumo de contenido de TV con amigos o familiares separados geográficamente (es decir, en hogares/lugares diferentes). Algunas respuestas cuyo valor mostrado en las gráficas es del 100% (o valores cercanos) de los participantes, como, por ejemplo, en el caso de la respuesta "Sí, siempre" de la Fig. 124, no se consideran relevantes, puesto que dichas gráficas se corresponden con porcentajes bajos de respuestas (un 2% de los participantes en el caso del ejemplo, equivalente, para el número total de muestras recogidas, a una única persona, que en ese caso ha sido un hombre).

EVALUACIÓN DE LA SINCRONIZACIÓN INTER-DESTINATARIO (+ INTER-DISPOSITIVO)

Los participantes también fueron preguntados acerca de parámetros y características del escenario como, por ejemplo, el comportamiento de los dispositivos secundarios con la herramienta de chat, la precisión (percibida) del nivel de sincronización alcanzado, la interfaz de usuario, etc. Las respuestas obtenidas se resumen a continuación

Respecto a la herramienta de chat desarrollada para la plataforma y adaptada al escenario de IDMS bajo evaluación, más de 66% de los participantes afirmaron preferir comunicarse a través de un chat privado, sincronizado e interactivo en lugar de utilizar uno público externo, como los de redes sociales populares (Fig. 127). Esto se debió a diversos factores: i) la existencia de un mecanismo de sincronización que evita *spoilers*; ii) la privacidad en las sesiones, a diferencia de cuando se usan las redes sociales; y iii) la inclusión de mensajes contextualizados predefinidos, que permiten comunicaciones más rápidas e interactivas, además de experiencias personalizadas y que les 'enganchan' más. De hecho, el 74% de los participantes opinó que este tipo de interacción con mensajes contextualizados ya predefinidos, enviados con simplemente pulsar un botón de la aplicación, permiten tener conversaciones más fluidas durante la sesión (Fig. 128).

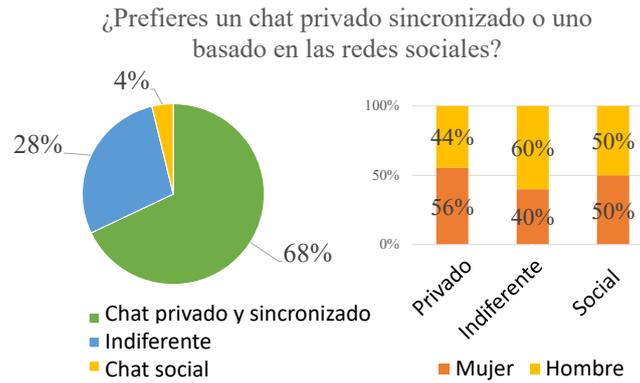


Fig. 127 Preferencias respecto a comunicarse a través de un chat privado sincronizado o público sin sincronizar (a través de redes sociales)

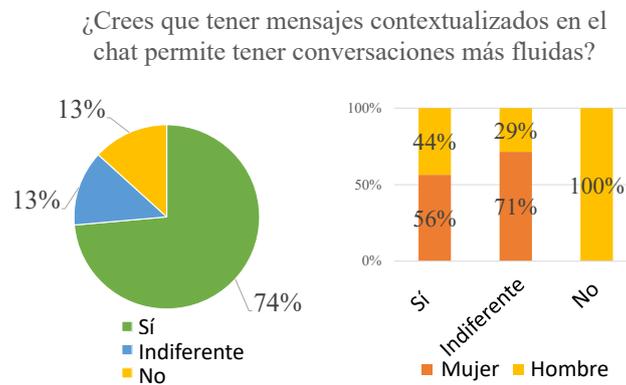


Fig. 128 Opinión respecto a la utilidad de tener mensajes predefinidos contextualizados

Respecto a la sincronización IDMS alcanzada, los participantes no tuvieron la sensación de estar en puntos de reproducción del contenido adelantados o retrasados (76% y 82%, respectivamente) con respecto a los del resto de usuarios involucrados en la sesión (Fig. 129), y tampoco consideraron haber recibido mensajes de texto a destiempo respecto al contenido que estaban visualizando (Fig. 130).



Fig. 129 Percepción del nivel de sincronización alcanzado



Fig. 130 Percepción del funcionamiento del mecanismo de sincronización de la herramienta de chat

Cabe resaltar que, cuando se preguntó a los participantes sobre el origen o la causa de haber recibido algún mensaje retrasado respecto al contenido visualizado, el 64.2% afirmó que esto se debía a un factor humano (por ejemplo, por ser relativamente lentos al escribir los mensajes). Además, el 84.9% de los participantes han declarado que los retardos experimentados en los mensajes de chat han sido razonables y no han llegado a ser molestos en ninguna ocasión.

Los resultados obtenidos en esta evaluación subjetiva corroboran el comportamiento satisfactorio del mecanismo de sincronización IDMS, pues son coherentes con los obtenidos durante la evaluación objetiva, presentados en el capítulo anterior en la sección correspondiente al escenario IDMS. Además, en esta fase de la evaluación tampoco se han encontrado diferencias significativas al sesgar por género los resultados obtenidos.

USABILIDAD, INTERÉS DESPERTADO Y APLICABILIDAD DEL ESCENARIO

Al finalizar la experiencia de uso libre del escenario, los participantes aportaron su opinión respecto a su utilidad y su impresión respecto al potencial del mismo. Tal y como se puede observar en la Fig. 131, un 89% de los participantes declararon que este escenario mejora la sensación de compañía (en inglés, *network togetherness*), un 95% declaró que el escenario es fácil de usar (Fig. 132) y un 94% que se trataba de un escenario útil e interesante (Fig. 133).

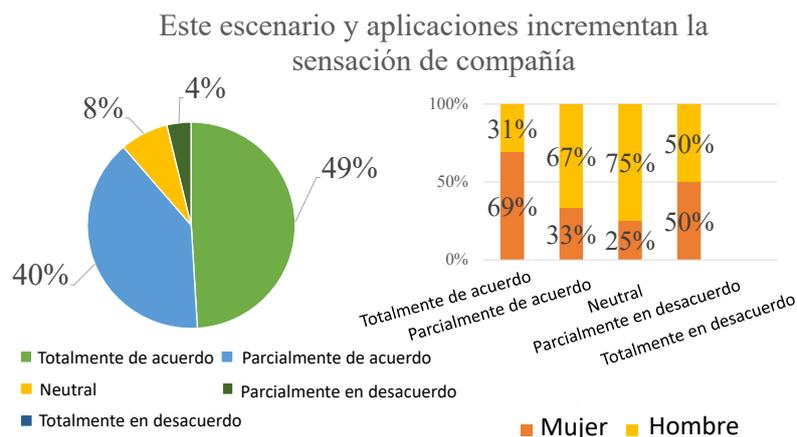


Fig. 131 Sensación de compañía durante la experiencia de uso del escenario

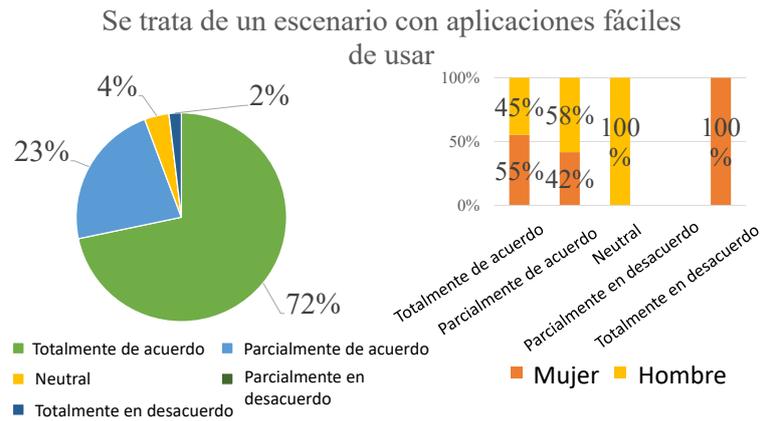


Fig. 132 Usabilidad del escenario

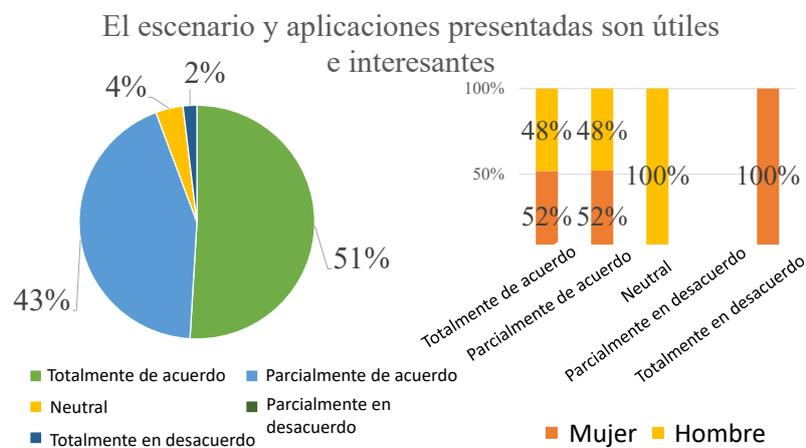


Fig. 133 Interés y utilidad del escenario

Por un lado, el 74% de los participantes consideró que el escenario evaluado solucionaba un problema real (Fig. 134). Por otro lado, un 78% aseguró que el escenario evaluado de TV Social podría tener un impacto en el mercado actual de consumo de TV (Fig. 135). Además, un 83% de los participantes lo utilizaría si estuviera disponible comercialmente (Fig. 136).

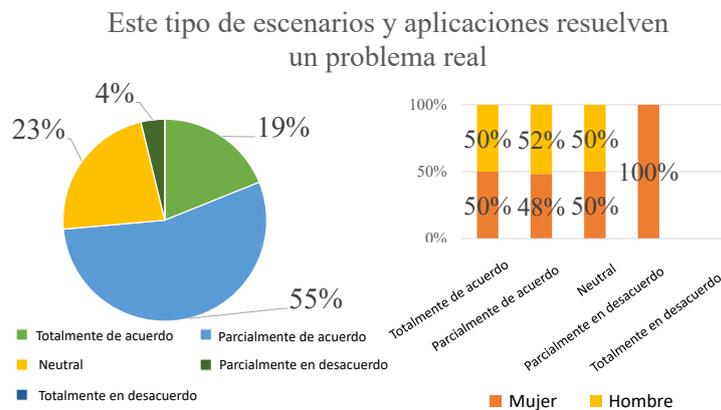


Fig. 134 Respuesta respecto a la utilidad real del escenario

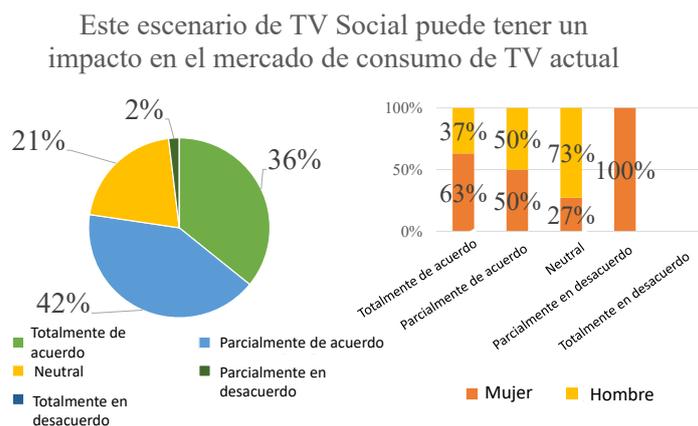


Fig. 135 Posible impacto en el mercado actual según los participantes

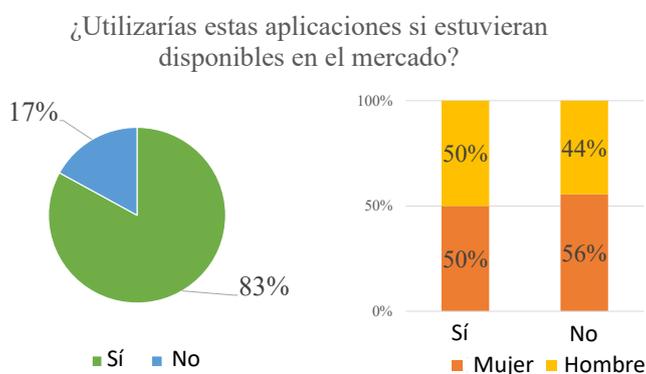


Fig. 136 Uso de las aplicaciones si estuvieran disponibles

Respecto a la opinión acerca de si el escenario implementado en la plataforma desarrollada en la Tesis Doctoral proporciona o facilita una experiencia de consumo más inmersiva, enriquecida y personalizable, el 96% estaban de acuerdo con esta afirmación (Fig. 137).

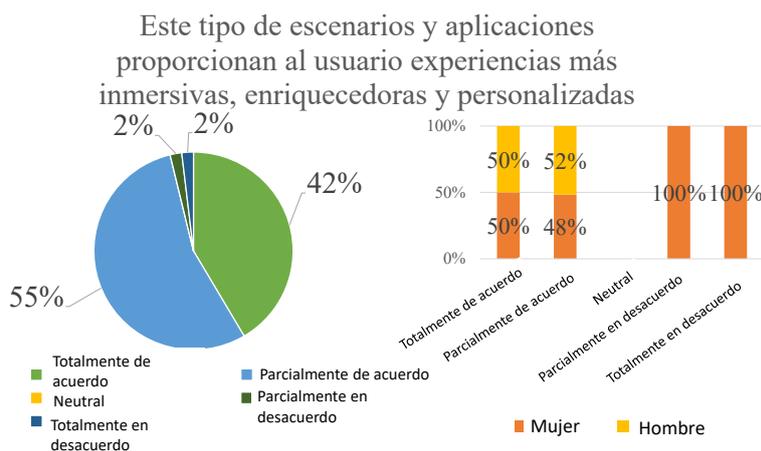


Fig. 137 Opinión sobre la capacidad de proporcionar experiencias de consumo más inmersivas, enriquecidas y personalizables

Además, un 87% de los participantes consideró que este tipo de escenarios también se podrían aplicar en otras áreas más allá del TV Social como, por ejemplo, para aplicaciones o servicios de aprendizaje a distancia (e-learning), lo cual es una muestra de su amplia aplicabilidad (Fig. 138).

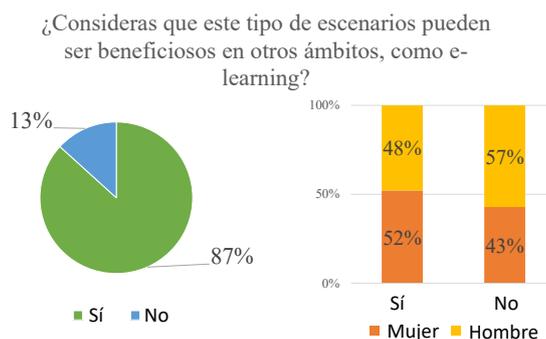


Fig. 138 Aplicabilidad de la plataforma en otras áreas

Finalmente, tal y como se puede observar desde la Fig. 134 hasta la Fig. 138, al sesgar los resultados por género, las diferencias no son relevantes, puesto que se puede observar que las opiniones son similares independientemente de quién las responda.

7.5.5 DISCUSIÓN

Tras observar los resultados obtenidos en esta evaluación subjetiva, se puede concluir que: 1) la solución propuesta soluciona un problema real que, además, ha sido experimentado en plataformas similares por alguno de los participantes, puesto que en la actualidad existe una falta de sincronización entre usuarios remotos compartiendo experiencias de consumo de contenidos multimedia que debe corregirse; 2) la QoE percibida por los participantes se ve afectada negativamente si el nivel de sincronización empeora, lo cual justifica la utilidad y la necesidad de establecer mecanismos de sincronización IDMS; 3) los participantes han reflejado en sus respuestas un interés en este tipo de escenarios, ya que consideran que les permiten compartir el consumo de contenido multimedia de forma enriquecida, y más personalizada e inmersiva.

7.6 ESCENARIO MULSEMEDIA

En esta sección se va a evaluar la QoE percibida por los participantes en un escenario mulsemmedia, con IDES e incluyendo contenido omnidireccional. El contenido principal se reproduce en el dispositivo principal (ejecutando la MS app), la TV, mientras que el contenido omnidireccional (secundario) se reproduce en un smartphone Samsung Galaxy Note 9 integrado en un HMD

Respecto a los contenidos audiovisuales disponibles, el contenido de TV principal que se visualiza en el dispositivo principal, que es transmitido a través de la red broadcast, se ha extraído previamente de contenido omnidireccional original (ver Tabla 35). En dicha tabla se incluye la información necesaria para su correcta reproducción. Como resumen, el contenido empleado para el escenario mulsemmedia se muestra en la Tabla 40.

Tabla 40 Contenido generado para la evaluación subjetiva

Título	Duración (s)	Codificación	Vistas 2D	Vistas 360	Tipo de proyección	Fuente
F1 & Roller Coaster	360	H.264 + AAC	3	1	CMP	https://www.youtube.com/watch?v=fQoVFraBOnC&t=9s https://www.youtube.com/watch?v=8lsB-P8nGSM&t=89s

Adicionalmente, los efectos multisensoriales incluidos en el escenario se presentan en la Tabla 41.

Tabla 41 Efectos multisensoriales incluidos en el escenario de evaluación

Tipo de efecto	Comentarios
Luz ambiental	Se trata de un efecto que proyecta los colores del contenido audiovisual sobre la pared en la que se encuentra la TV. Ha estado activo durante todo el consumo de contenido.
Aromas	Se han empleado dos aromas diferentes, uno de neumático quemado (para contenido F1) y otro de pino (para contenido Roller Coaster). Se ha activado en momentos puntuales para dar énfasis en determinadas escenas.
Viento	Se ha activado en momentos específicos para dar una sensación de velocidad inmersiva.
Vibración	Se ha activado en momentos específicos para dar una sensación de movimiento inmersiva.

El espacio en el que se ha realizado la evaluación se muestra en la Fig. 139 y consta de un habitáculo de 9m² que incluye: una TV, un HMD, un sofá con un generador del efecto de vibración, un generador de aromas, un generador del efecto de viento y un generador de luz ambiental.



Fig. 139 Aspecto del escenario de evaluación subjetiva

7.6.1 PARTICIPANTES

El total de participantes para esta evaluación ha sido de 32 personas, de los cuales se proporciona información en la Tabla 36 (recogida durante la primera fase de la evaluación).

Tabla 42 Perfil de los participantes

Género		Edad		Nivel de estudios	
Hombre	Mujer	<21	3%	Secundaria	13%
55%	45%	21-25	19%	F.P.	25%
		26-30	13%	Grado	28%
		31-35	13%	Máster	19%
		36-40	0%	PhD	15%
		41-45	3%		
		46-50	28%		
		>50	21%		

Además de la información contenida en la tabla anterior, un 53% de los participantes cuenta con un perfil de trabajo técnico.

Cabe destacar que ninguno de los participantes presentaba deficiencias del tipo audiovisual, como, por ejemplo, problemas graves de visión o auditivos. Además, ninguno de ellos ha presentado ningún tipo de síntoma post-Covid (por ejemplo, falta de olfato) o congestión nasal que pudiera influir o afectar en la evaluación (sobre todo en los aspectos relacionados con el efecto de los aromas).

Respecto a hábitos de consumo de tabaco, un 15% de los participantes son fumadores activos (más de 1 cigarrillo diario), un 9% son fumadores ocasionales (menos de 1 cigarrillo diario), un 3% son fumadores pasivos y un 73% no son fumadores ni activos ni pasivos.

Siguiendo las recomendaciones presentadas en [106], se ha indicado a los participantes que no fumaran ni consumieran café o alimentos en un periodo previo a la evaluación de 2h.

7.6.2 METODOLOGÍA DE EVALUACIÓN

La evaluación de este escenario se ha llevado a cabo por parejas y se ha dividido en las fases definidas al inicio del capítulo.

Durante la fase 0 de la evaluación, los participantes han recibido una breve introducción a la plataforma, y se les ha explicado sus principales características, así como los del tipo de contenido que se iba a consumir.

Para la evaluación de este escenario, los participantes han llevado a cabo un proceso de familiarización durante la fase 1, en la que el consumo de contenido ha sido exclusivamente audiovisual, es decir, sin ningún tipo de efecto multisensorial activado. La motivación de no tener ningún efecto multisensorial activado ha radicado en tratar de evidenciar una posible variación en la QoE percibida por los participantes dependiendo de si este tipo de efectos están activos o no.

A continuación, durante la fase 2, se han activado los cuatro efectos multisensoriales indicados. En particular, durante la primera mitad de la fase 2, uno de los dos participantes de cada sesión ha consumido el contenido a través de la TV mientras que el otro participante lo ha hecho a través del HMD. En la segunda mitad de la fase 2, se han intercambiado los dispositivos de consumo con el fin de que ambos participantes hayan tenido un consumo equivalente del contenido tanto

a través de la TV como del HMD. Finalmente, durante la fase 3 se ha recogido las opiniones posteriores a la experiencia de consumo.

7.6.3 RESULTADOS OBTENIDOS

A continuación, se muestran los resultados obtenidos para este escenario en particular.

HÁBITOS Y EXPERIENCIA PREVIA

Respecto a la experiencia previa de los participantes con contenidos multisensoriales, los resultados se muestran en la Fig. 140. Cabe destacar que un 41% de los participantes respondieron no haber tenido ninguna experiencia previa con este tipo de efectos y un 38% había experimentado con contenidos multisensoriales en parques temáticos o atracciones.

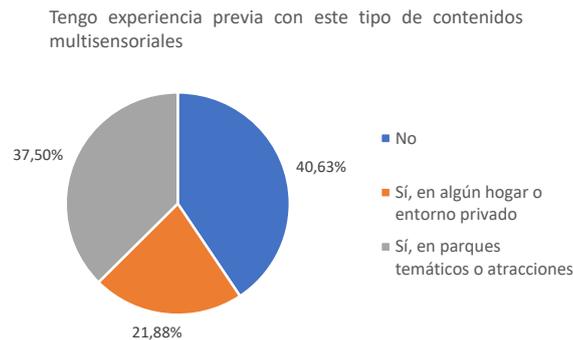


Fig. 140 Experiencia previa con contenidos multisensoriales

Por otra parte, respecto a la opinión preliminar acerca de si los efectos multisensoriales mejoran la experiencia de usuario, el 75% de los participantes afirmó estar totalmente de acuerdo con dicha afirmación, un 22% parcialmente de acuerdo y solamente un 3% afirmó estar parcialmente en desacuerdo.

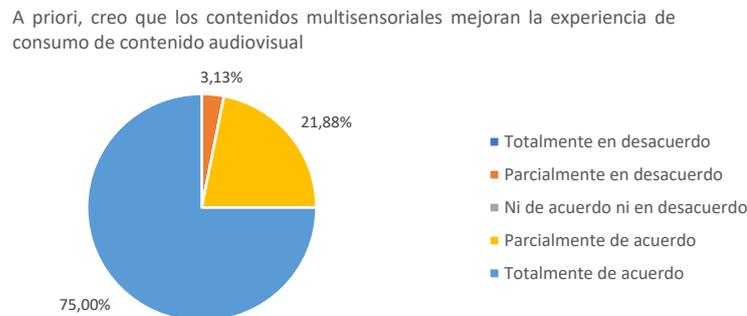


Fig. 141 Opinión preliminar acerca de la mejora de la experiencia de consumo con la inclusión de efectos multisensoriales

PERCEPCIÓN DE LA QoE Y DE LA SINCRONIZACIÓN DE EFECTOS SENSORIALES

Tras haber completado las fases 1 y 2, los participantes evaluaron la experiencia de consumo de contenido multimedia a través de TV y del HMD, obteniendo unos valores medios de MOS de 4.03 y 4.19 con unos I.C. 95% de 0.27 y 0.26, respectivamente. En particular, durante el consumo de contenido a través de la TV, un 97% de los participantes consideró que "los efectos multisensoriales estaban sincronizados con el contenido audiovisual"; el 3% restante se mostró indeciso respecto a dicha afirmación (ver Fig. 142).

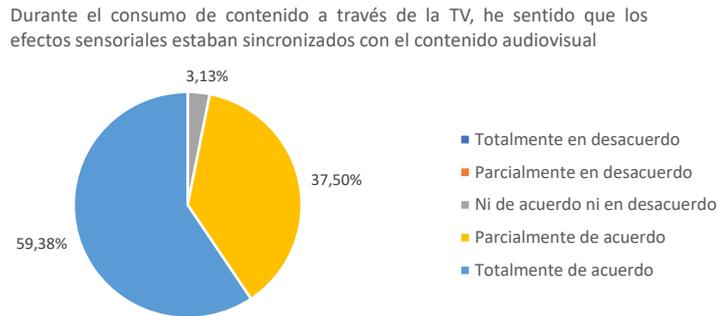


Fig. 142 Percepción de sincronización entre contenidos presentados en la TV y los efectos multisensoriales

Sin embargo, durante el consumo de contenido a través del HMD, la totalidad de los participantes coincidieron en afirmar que “los efectos multisensoriales estaban sincronizados con el contenido audiovisual” (ver Fig. 143).

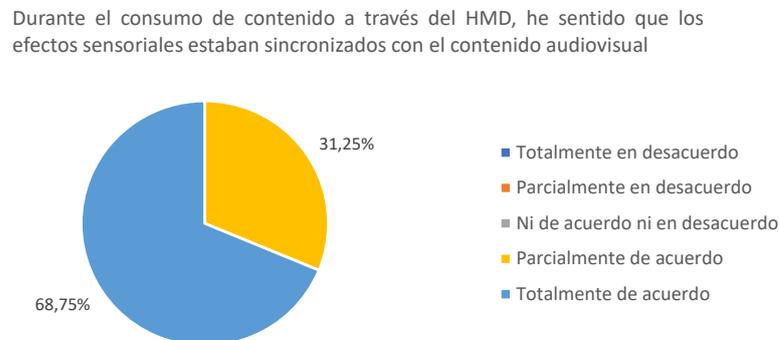


Fig. 143 Percepción de sincronización entre contenidos presentados en el HMD y los efectos multisensoriales

Respecto a la interacción entre los participantes, se ha preguntó a los que consumieron contenido a través de la TV si consideraban que su compañero con HMD estaba consumiendo el mismo contenido de forma sincronizada, estando el 75% de los participantes de acuerdo con esta afirmación y un 25% ni de acuerdo ni en desacuerdo.

Respecto a la preferencia por algún efecto multisensorial determinado, la respuesta de los participantes se muestra en la Fig. 144. Se puede observar que el efecto de vibración resultó ser el preferido por los participantes, seguido por el viento, los aromas y finalmente, la luz ambiental. Por otro lado, en la Fig. 145 se muestran los efectos que resultaron más molestos para los participantes. Un 16% de los participantes consideró molesto el efecto de viento, un 6% señaló la vibración y un 3% los aromas. El 75% restante no consideró molesto ninguno de los efectos multisensoriales incluidos en el escenario.

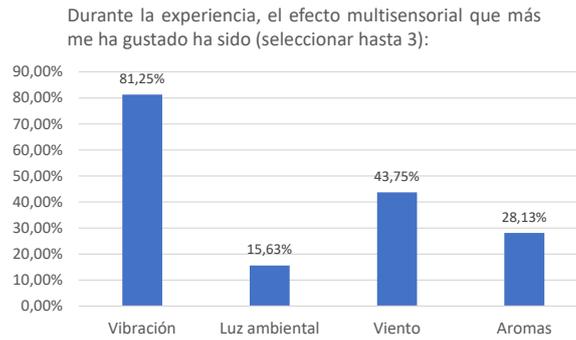


Fig. 144 Preferencia de los usuarios por algún efecto multisensorial

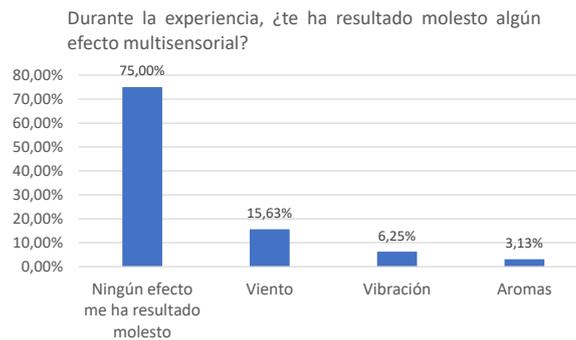


Fig. 145 Rechazo de los participantes hacia alguno de los efectos multisensoriales percibidos

Finalmente, cuando se les preguntó a los participantes si consideran que los efectos sensoriales incluidos mejoraron su experiencia personal durante el consumo de contenidos, un 91% consideró que así fue y el 9% restante, consideró que no ha afectó ni positiva ni negativamente en la experiencia de consumo.

INTERÉS DESPERTADO Y APLICABILIDAD DEL ESCENARIO

Respecto al interés despertado en los usuarios, la totalidad de los participantes consideran que, en general, los efectos multisensoriales mejoraron la experiencia de consumo de contenido audiovisual (ver Fig. 146).

Creo que los contenidos multisensoriales mejoran la experiencia de consumo de contenido audiovisual.

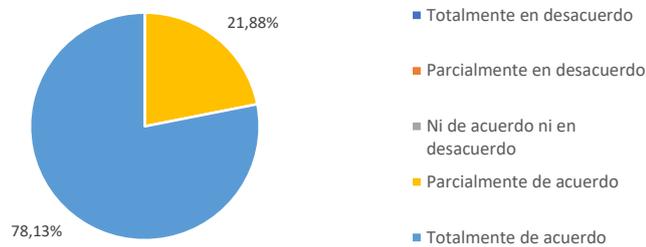


Fig. 146 Opinión de los participantes acerca de la mejora de la QoE percibida si hay efectos multisensoriales presentes

Adicionalmente, en las experiencias de consumo multisensorial, el 72% de los participantes prefirió utilizar el HMD y el 28% restante la TV para el consumo de contenido audiovisual.

A continuación, se les preguntó si preferirían contar con un sistema de consumo mulsemmedia en el hogar formado por una TV (ver Fig. 147), o bien por una TV y dispositivos HMD (ver Fig. 148).

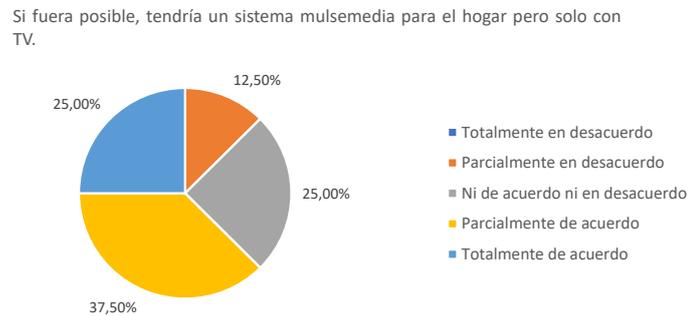


Fig. 147 Interés de los participantes en poseer un sistema mulsemmedia (solo TV)

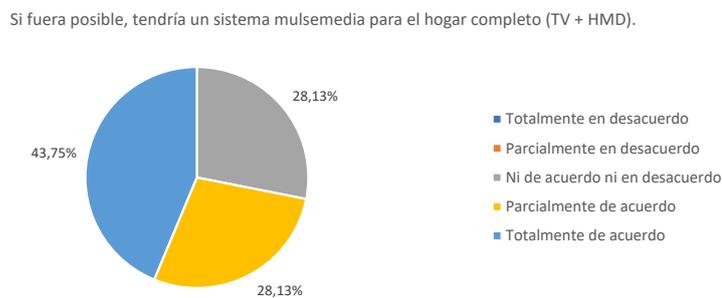


Fig. 148 Interés de los participantes en poseer un sistema mulsemmedia (TV + HMD)

Se puede observar cómo existe un mayor número de participantes que prefirieron contar con ambos tipos de dispositivo de consumo (TV y HMD).

Adicionalmente, respecto a su disponibilidad para ir a locales o eventos en los que se ofreciera este tipo de sistemas de consumo de contenido multisensorial, los resultados se muestran en la Fig. 149. Un 18.75% de los participantes iría siempre y un 25% casi siempre. La respuesta más seleccionada fue "de vez en cuando" con un 34.38%; y un 21.88% de los participantes iría en muy contadas ocasiones.

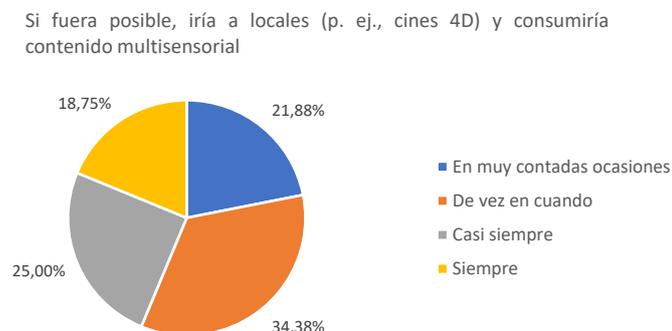


Fig. 149 Predisposición de los participantes para ir a lugares o eventos donde se ofrezca experiencias de consumo mulsemmedia

Finalmente, cuando los participantes fueron preguntados sobre qué expectativas tienen para poder contar con este tipo de sistemas en el hogar en el futuro, un 53% considera que entre 2 a

5 años vista; un 25% espera contar con este tipo de sistemas en menos de 2 años; un 16% entre 5 a 10 años vista; y finalmente un 6% a más de 10 años vista.

7.6.4 DISCUSIÓN

Tras analizar los resultados extraídos de la evaluación subjetiva, se puede deducir que el nivel de sincronización percibido entre el contenido audiovisual y los efectos multisensoriales mejora cuando el dispositivo de consumo utilizado es el HMD y que este tipo de dispositivos inmersivos mejora la experiencia de consumo al influir en el participante a la hora de tolerar valores de sincronización menos favorables.

Respecto a los diferentes tipos de efectos multisensoriales que han experimentado los participantes, hay una clara preferencia por el efecto de vibración y un interés significativamente más bajo por la iluminación ambiental. Esto puede deberse por la sensación de novedad, puesto que el primero no es tan común entre los hogares, pero en cambio, el efecto de luz ambiental ya está ampliamente integrado en los hogares a través de sistemas como el Ambilight, integrado en muchos modelos de TV del fabricante Philips. Respecto al efecto de viento, algunos participantes lo han considerado molesto, por lo que se evidencia que se requiere profundizar en este tipo de efectos para llegar a un compromiso entre una percepción evidente del propio efecto sin que afecte de forma negativa a la experiencia de consumo.

Por otro lado, de los resultados se puede considerar que, a pesar de ser de interés para los participantes, este tipo de sistemas de consumo de contenido multimedial no cuenta con una expectativa de estar disponible a corto plazo. Sin embargo, cuando comparamos los resultados previos y posteriores a la experiencia de consumo, la opinión de los participantes mejora cuando se les pregunta si consideran que este tipo de experiencias mejoran la QoE percibida (ver Fig. 141 y Fig. 146).

7.7 CONCLUSIONES

En este capítulo se han presentado los resultados obtenidos para las evaluaciones subjetivas de cuatro escenarios diferentes: IDES básico, IDES inmersivo (contenido omnidireccional), IDMS (+ IDES) e IDES multimedial. En todos los escenarios se observan resultados muy positivos y satisfactorios respecto a las experiencias de consumo llevadas a cabo por los participantes, así como una percepción de un nivel de sincronización suficiente para garantizar una buena QoE a los usuarios. Esto ha permitido despertar el interés de los participantes para cada uno de los escenarios experimentados, lo cual evidencia un previsible impacto significativo en el mercado de consumo de contenidos audiovisuales actual, puesto que las tecnologías adoptadas para diseñar las soluciones propuestas en la Tesis Doctoral, para alcanzar un nivel de sincronización satisfactorio ya se encuentran plenamente integradas en el mercado.

Sin embargo, para tal fin se debe considerar el uso de plataformas como la presentada en esta Tesis Doctoral, cuyas funcionalidades son esenciales puesto que permiten alcanzar un nivel de sincronización satisfactorio para el usuario, a la vez que le proporciona una serie de herramientas y funcionalidades que facilitan la usabilidad y la interacción con las aplicaciones integradas en la plataforma.

Por tanto, en lo que respecta a la evaluación de la QoE percibida por los participantes, los resultados obtenidos para estos cuatro escenarios permiten afirmar que el funcionamiento tanto de los mecanismos de sincronización desarrollados en esta Tesis Doctoral, como de la plataforma desarrollada para evaluarlos, es el adecuado. Permite proporcionar una buena QoE de consumo de contenidos de TV híbrida, incluyendo contenidos omnidireccionales y efectos multisensoriales,

tanto en experiencias individuales como compartidas con otros usuarios remotos, y tanto utilizando un dispositivo (terminal híbrido) como varios dispositivos (principal y secundarios).

Capítulo 8. CONCLUSIONES Y TRABAJO FUTURO

En este capítulo se presentan las conclusiones y las contribuciones de esta Tesis Doctoral. Adicionalmente, se expone el trabajo futuro derivado la misma.

8.1 CONCLUSIONES

Tradicionalmente, los servicios multimedia estaban concebidos para ser transmitidos a través de redes específicas y consumidos en dispositivos concretos. Un claro ejemplo de esto es la televisión tradicional, cuyos contenidos se transmiten a través de la red broadcast y se consumen a través de un televisor convencional.

Hoy en día, existe una gran variedad de tecnologías de entrega de contenidos, dispositivos de consumo y contenidos multimedia a disposición de los consumidores. Respecto a la entrega de contenidos multimedia, ésta puede realizarse a través de tecnologías broadcast o broadband. Por otro lado, las tecnologías broadband proporcionan servicios bidireccionales, interactivos y adaptativos, según los recursos disponibles o las preferencias de los consumidores.

Respecto a los dispositivos de consumo multimedia, sin tener en cuenta la amplia variedad de TVs conectadas, la proliferación y utilización de diferentes dispositivos secundarios para dicho fin (cmoo, por ejemplo, portátiles, tablets o smartphones) es ya una realidad. Por una parte, estos dispositivos cuentan con múltiples tecnologías de conexión (por ejemplo, WiFi, 4G...), por lo que son capaces de recibir contenidos multimedia a través de todas ellas. Por otra parte, dichos dispositivos también cuentan con una gran heterogeneidad de recursos, rendimiento y capacidades, sobre todo en términos de disponibilidad de ancho de banda, soporte a diferentes tipos de tecnología o a tipos de contenido multimedia, capacidad de procesamiento y tamaño de las pantallas.

Por tanto, se puede observar que se está produciendo una evolución con relación a la forma de consumir contenido, y dicha evolución permite mejorar la QoE del consumidor, aumentando la disponibilidad de dispositivos de consumo y proporcionando unos servicios multimedia más adaptativos y omnipresentes.

Resulta evidente que, cuando dos o más dispositivos están involucrados en la experiencia de consumo, también es necesario establecer mecanismos de sincronización entre los procesos de reproducción de todos ellos (IDMS, si involucra dispositivos remotos, o IDES, en el caso de involucrar dispositivos locales) con el objetivo de proporcionar una experiencia de consumo satisfactoria.

Así como la oferta de contenido a través de redes broadband ha experimentado un gran auge, recientemente la experiencia de consumo de contenido ha dejado de involucrar exclusivamente a los sentidos de la vista y el oído. Con el fin de estimular el resto de los sentidos, ha surgido el nuevo concepto de contenido mulsemmedia. Dicho contenido permite proporcionar al usuario estímulos para otros sentidos, además de los dos citados, a través de otro tipo de dispositivos

adicionales que permiten la generación de efectos sensoriales, como, por ejemplo, viento, calor, frío, luz ambiental, niebla, vibración, aromas, etc. En definitiva, el principal objetivo de incluir este nuevo tipo de contenido en las experiencias de consumo multimedia es proporcionar al usuario una experiencia más inmersiva y realista ("*Ver es creer, pero sentir es la realidad*" -T. Fuller [16]).

Por tanto, se puede afirmar que el nuevo paradigma de experiencia de consumo multimedia que se presenta cuenta con las tecnologías y dispositivos necesarios para poder ser social, inmersivo y multisensorial.

Para asegurar una buena QoE a los usuarios, resulta necesario, entre otros desarrollos complementarios, el diseño y establecimiento de mecanismos de sincronización híbridos que sean agnósticos e independientes de la tecnología de distribución de los contenidos. En esta Tesis Doctoral, se han estudiado los diferentes mecanismos de sincronización híbridos, y se ha centrado el foco en el caso de uso de *TV Social*, diseñando una solución de sincronización híbrida para escenarios IDMS e IDDES, generando contenido multisensorial y validando dicha solución en una plataforma híbrida compatible con el estándar HbbTV. Se ha abordado, principalmente, el diseño, implementación y validación de mecanismos de sincronización híbridos, teniendo en cuenta la naturaleza del contenido (multimedia y/o mulsemmedia) y su entorno de consumo (multidispositivo y/o multidestinario), validando posteriormente la calidad de la experiencia percibida por parte de los consumidores (QoE).

Primeramente, en el Capítulo 2 se ha llevado a cabo una revisión de los conceptos, estándares y mecanismos que permiten el establecimiento de mecanismos de generación, transmisión y consumo sincronizado e inmersivo de contenido multimedia híbrido y multisensorial. Por un lado, se ha realizado una revisión de los estándares y middlewares relacionados con la transmisión y consumo de contenido híbrido. Por otro lado, se han presentado los diferentes tipos de sincronización, las estrategias de sincronización y los esquemas de control que pueden ser adoptados para alcanzar un nivel de sincronización satisfactorio. Además, se ha expuesto las características más significativas de contenidos de nueva generación (contenidos omnidireccional y multisensorial) contemplados en la Tesis Doctoral, así como las técnicas actuales para su codificación, transmisión y consumo.

En el Capítulo 3, se ha presentado y descrito trabajos relacionados en el área de investigación de esta Tesis Doctoral, tanto en trabajos relacionados con la sincronización híbrida de contenidos, tanto en uno o varios dispositivos (sincronización híbrida IDDES), como en diferentes destinatarios (sincronización híbrida IDMS). Adicionalmente, se han expuesto trabajos previos relacionados con contenidos de nueva generación (como, por ejemplo, vídeo omnidireccional o contenido mulsemmedia).

Se han descrito con detalle el conjunto de mecanismos de sincronización híbrida propuestos para posibilitar nuevas experiencias televisivas enriquecidas, personalizadas, inmersivas y compartidas (Capítulo 4 de la Tesis doctoral). Previamente, se han descrito los requisitos necesarios para que se puedan establecer satisfactoriamente estos mecanismos de sincronización, como: 1) los contenidos híbridos deben estar relacionados de alguna manera, con el fin, por ejemplo, de poder localizar contenido complementario al contenido principal; 2) los procesos, dispositivos o destinatarios (según escenario) deben descubrirse y deben establecer algún tipo de canal bidireccional que les permita comunicarse entre sí; 3) los procesos, dispositivos o destinatarios involucrados en la misma sesión de consumo deben soportar y reproducir contenido recibido a través de la red broadcast y también a través de la red broadband; y 4) los procesos de reproducción involucrados deben soportar mecanismos de ajuste del estado de reproducción con el fin de alcanzar un nivel de sincronización satisfactorio.

Respecto a los escenarios mono- y multidispositivo (IDDES), se han expuesto las siguientes contribuciones: i) especificación y diseño de necesidades clave para alcanzar un nivel de

sincronización IDES satisfactorio, como son la señalización de contenido relacionado; ii) la interacción y coordinación entre dispositivos principales y secundarios; y iii) mecanismos específicos de sincronización. Respecto a la solución de sincronización IDMS que se ha propuesto en esta Tesis Doctoral, se han aportado las siguientes contribuciones: i) especificación y diseño de necesidades clave para alcanzar un nivel de sincronización IDMS satisfactorio, como son la señalización sesiones de consumo de contenido audiovisual disponibles; ii) la interacción y coordinación entre dispositivos principales en cada destinatario; iii) mecanismos específicos de sincronización; y iv) canales de comunicación sincronizados entre los usuarios de una misma sesión de consumo de contenido audiovisual.

En particular, esta Tesis Doctoral se ha centrado, principalmente, en proporcionar unos niveles de sincronización IDES/IDMS satisfactorios, así como unos canales de interacción entre usuarios de una sesión de consumo compartida.

Respecto a la generación y consumo de contenido mulsemmedia, las principales contribuciones en esta Tesis Doctoral son: i) el diseño e implementación de una solución capaz de generar, codificar (los metadatos asociados) y presentar contenido mulsemmedia relacionado con contenidos recibidos a través de redes broadcast y broadband; ii) un mecanismo que permite señalar la existencia de efectos sensoriales de forma integrada dentro del flujo de transporte MPEG2-TS; iii) el diseño de un controlador de efectos multisensoriales, capaz de analizar y dar soporte a cualquier efecto sensorial (siempre y cuando se disponga de los materiales y dispositivos encargados de generar dichos efectos); iv) diseño de una interfaz (API) para la comunicación entre los dispositivos de consumo y el controlador mulsemmedia; v) diseño e implementación de un mecanismo IDES para sincronizar la reproducción del contenido audiovisual y la generación de los efectos sensoriales; y vi) la integración del soporte para contenido mulsemmedia en la plataforma extremo-a-extremo del grupo de I+D en el que se enmarca esta Tesis Doctoral, adaptada para los casos de uso de los escenarios de evaluación.

Se ha implementado una plataforma extremo-a-extremo de generación y consumo de contenidos híbridos, multidestinataro, multidispositivo y multisensoriales, en la que se han incluido todas las soluciones propuestas en la Tesis Doctoral (Capítulo 5). Se ha detallado el desarrollo de su arquitectura y los módulos incluidos en la misma. También se ha descrito cada una de las funcionalidades incluidas en la misma: la generación del contenido multimedia, la inserción de los descriptores TEMI location y timestamp, la generación de etiquetas mulsemmedia, la señalización de todo el contenido complementario disponible a través de la red broadband, así como toda la parte correspondiente a la recepción de los contenidos y la interacción entre los dispositivos de consumo en el extremo del usuario final.

Se han integrado en la plataforma los mecanismos de sincronización presentados en esta Tesis Doctoral, con el fin de evaluarlos, tanto objetiva como subjetivamente, en diferentes casos de uso, tal y como se expone en los dos capítulos siguientes.

Las soluciones propuestas han sido evaluadas, tanto objetiva como subjetivamente, obteniendo resultados más que satisfactorios (Capítulos 6 y 7). Se han descrito los escenarios, la metodología seguida y los resultados de las evaluaciones objetivas llevadas a cabo durante la Tesis Doctoral, con respecto al nivel de sincronización híbrida IDES e IDMS alcanzado (Capítulo 6). Se ha evaluado la sincronización en varios escenarios: un primer escenario básico incluyendo una solución IDES (sin contenido omnidireccional), el cual incluye contenido multivista; un segundo escenario inmersivo incluyendo una solución IDES con contenido omnidireccional; y finalmente, un tercer escenario incluyendo soluciones de sincronización híbrida IDES e IDMS.

En los escenarios de evaluación IDES, se ha demostrado que, aun habiendo pérdidas de paquetes en la transmisión broadcast, y retardos en la transmisión broadband, el nivel de

sincronización entre las MS y CS apps se mantienen siempre por debajo del umbral de asincronía máxima configurado.

Por otro lado, en el escenario IDMS también se ha corroborado que los valores de asincronía registrados entre destinatarios se mantienen por debajo del umbral de la asincronía de la sesión máxima configurado. Respecto a la combinación de soluciones de sincronización IDMS + IDES, se observa cómo los ajustes IDMS en una MS app afectan de forma directa a los ajustes de la CS app enlazada, pudiendo afectar negativamente en la QoE percibida por los usuarios en momentos puntuales.

Se han realizado evaluaciones subjetivas (Capítulo 7) en cuatro escenarios diferentes: IDES básico, IDES inmersivo (contenido omnidireccional), IDMS (+ IDES) e IDES mulsemmedia. En todos los escenarios se han observado resultados muy positivos y satisfactorios respecto a las experiencias de consumo llevadas a cabo por los participantes, así como una percepción de un nivel de sincronización suficiente para garantizar una buena QoE a los usuarios. Esto ha permitido despertar el interés de los participantes para cada uno de los escenarios experimentados, lo cual evidencia un previsible impacto significativo en el mercado de consumo de contenidos audiovisuales actual, puesto que las tecnologías adoptadas para diseñar las soluciones propuestas en la Tesis Doctoral, para alcanzar un nivel de sincronización satisfactorio ya se encuentran plenamente integradas en el mercado.

Por tanto, en lo que respecta a la evaluación de la QoE percibida por los participantes, los resultados obtenidos para estos cuatro escenarios permiten afirmar que el funcionamiento tanto de los mecanismos de sincronización desarrollados en esta Tesis Doctoral, como de la plataforma desarrollada para evaluarlos, es el adecuado. Permite proporcionar una buena QoE de consumo de contenidos de TV híbrida, incluyendo contenidos omnidireccionales y efectos multisensoriales, tanto en experiencias individuales como compartidas con otros usuarios remotos, y tanto utilizando un dispositivo (terminal híbrido) como varios dispositivos (principal y secundarios).

Finalmente, cabe destacar que durante el desarrollo de esta Tesis Doctoral, se han publicado 9 artículos en revistas especializadas (y se ha enviado otro que está actualmente en proceso de revisión); 8 artículos para congresos nacionales e internacionales; un capítulo de libro; y se han codirigido dos TFGs relacionados con la temática de investigación. Los datos de dichas publicaciones pueden consultarse en el Anexo *Publicaciones y trabajos derivados de la Tesis Doctoral*.

8.2 TRABAJO FUTURO

Como trabajo futuro, se pretende desarrollar una versión más actualizada de la plataforma, en la que se soporte contenido con codificaciones actuales, como HEVC. Ello permitirá adoptar estrategias de transmisión de contenido del tipo tiled-based streaming para contenido omnidireccional, adaptativas y dependientes del FoV del consumidor. Adicionalmente, para mejorar la compatibilidad con dispositivos de consumo que no sean compatibles con el protocolo HEVC, se llevará a cabo el diseño y desarrollo de procesado en la nube (edge-based), para poder entregar al consumidor contenido omnidireccional de forma más eficiente, en términos de ancho de banda y del uso de recursos de los dispositivos de consumo.

Respecto al contenido mulsemmedia, se analizará de forma objetiva el nivel de sincronización alcanzado, con el fin de proporcionar una solución más robusta y particularizada para el entorno donde se incluya este tipo de contenido.

Se pretende emplear tecnología 5G para observar el rendimiento de los mecanismos de sincronización propuestos en esta Tesis Doctoral, a través de las plataformas presentadas en la misma.

Además, se pretende adoptar otros mecanismos de monitorización y evaluación del nivel de sincronización alcanzado, como el presentado en [171], consistente en capturar y evaluar mediante técnicas de visión artificial el nivel de sincronización alcanzado mediante la comparación de fotogramas entre las distintas pantallas involucradas. También se buscará diseñar nuevos mecanismos de sincronización proactivos, que, basándose en técnicas de Inteligencia Artificial, se anticipen a la aparición de situaciones de asincronía y las eviten.

REFERENCIAS

- [1] A. Begen, T. Akgul, and M. Baugher, "Watching video over the web: Part 1: Streaming protocols," *IEEE Internet Comput.*, vol. 15, no. 2, pp. 54–63, Mar. 2011, doi: 10.1109/MIC.2010.155.
- [2] R. Frederick, "RTP: A Transport Protocol for Real-Time Applications," *Internet Soc. RFC 3550*, pp. 1–89, 2003.
- [3] R. Pantos and E. W. May, "Http Live Streaming. RFC 8216." 2017.
- [4] International Organization for Standardization, "ISO/IEC 23009-1:2014 Information technology -- Dynamic adaptive streaming over HTTP (DASH) -- Part 1: Media presentation description and segment formats," 2014.
- [5] HbbTV Association, "ETSI TS 102 796 V1.6.1 . HbbTV 2.0.3 Specification." 2020.
- [6] F. Boronat, M. Montagud, D. Marfil, and C. Luzon, "Hybrid Broadcast/Broadband TV Services and Media Synchronization: Demands, Preferences and Expectations of Spanish Consumers," *IEEE Trans. Broadcast.*, vol. 64, no. 1, pp. 52–69, 2018, doi: 10.1109/TBC.2017.2737819.
- [7] J. Le Feuvre, N. Viet-Thun-Trung, W. Hamidouche, M. Patrick, and D. Pascal, "A Test Bed for Hybrid Broadcast Broadband Services," p. 4, Jun. 2015.
- [8] R. Van Brandenburg and A. Veenhuizen, "Immersive second-screen experiences using hybrid media synchronization," in *Media Synchronization Workshop (MediaSync) 2013, NEM Summit 2013 workshop*, 2013.
- [9] M. Montagud, F. Boronat, H. Stokking, and R. Van Brandenburg, "Inter-destination multimedia synchronization: Schemes, use cases and standardization," *Multimed. Syst.*, vol. 18, no. 6, pp. 459–482, Nov. 2012, doi: 10.1007/s00530-012-0278-9.
- [10] We are social and Hootsuite, "Digital 2020," 2020. [Online]. Available: <https://wearesocial.com/digital-2020>.
- [11] Digital TV Europe, "Digital TV Europe Industry Survey 2019," 2019. [Online]. Available: <https://www.digitaltveurope.com/magazine/digital-tv-europe-industry-survey-2019/>.
- [12] YouGov, "YouGov: VR adoption grows to 11% of US adults – Digital TV Europe," 2019. [Online]. Available: <https://www.digitaltveurope.com/2019/03/22/yougov-vr-adoption-grows-to-11-of-us-adults/>.
- [13] Advanced Television, "Study: TV still the dominant screen for youngsters," 2019. [Online]. Available: <https://advanced-television.com/2019/02/19/study-more-screen-time-for-young-children/>.
- [14] Ebiquity, "Mind the Gap - A closer look at video advertising reach in the age of increasing media fragmentation," 2020. .
- [15] IAB Spain, "Estudio de Televisión Conectada 2019," 2019. [Online]. Available: <https://iabspain.es/estudio/estudio-de-television-conectada-2019/>.
- [16] G. Ghinea, C. Timmerer, W. Lin, and S. R. Gulliver, "Mulsemedia: State of the Art, Perspectives, and Challenges.," *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 11, no.

- 1s, pp. 1–23, Oct. 2014, doi: 10.1145/2617994.
- [17] M. O. Van Deventer, H. M. Stokking, O. A. Niamut, F. A. Walraven, and V. B. Klos, "Advanced interactive television services require content synchronization," in *Proceedings of IWSSIP 2008 - 15th International Conference on Systems, Signals and Image Processing*, 2008, pp. 109–112, doi: 10.1109/IWSSIP.2008.4604379.
- [18] M. Montagud, F. Boronat, H. Stokking, and R. Van Brandenburg, "Inter-destination multimedia synchronization: Schemes, use cases and standardization," *Multimed. Syst.*, vol. 18, no. 6, pp. 459–482, 2012, doi: 10.1007/s00530-012-0278-9.
- [19] W. J. Kooij, H. M. Stokking, R. Van Brandenburg, and P. T. De Boer, "Playout delay of tv signals: Measurement system design, validation and results," in *TVX 2014 - Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and Online Video*, 2014, pp. 23–30, doi: 10.1145/2602299.2602310.
- [20] L. Beloqui Yuste, F. Boronat, M. Montagud, and H. Melvin, "Understanding timelines within MPEG standards," *IEEE Commun. Surv. Tutorials*, vol. 18, no. 1, pp. 368–400, 2016, doi: 10.1109/COMST.2015.2488483.
- [21] D. Mills, U. Delaware, E. J. Martin, J. Burbank, and W. Kasch, "IETF Internet Standard RFC 5905. Network Time Protocol Version 4: Protocol and Algorithms Specification." 2010.
- [22] IEEE, "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems," *IEEE Std 1588-2002*, pp. i–144, 2002, doi: 10.1109/IEEESTD.2002.94144.
- [23] N. Borch, F. Daoust, and I. Arntzen, "Timing - Small step for developers, giant leap for the media industry," in *IBC 2016 Conference*, 2016, doi: 10.1049/ibc.2016.0013.
- [24] GStreamer, "GStreamer Framework." [Online]. Available: <https://gstreamer.freedesktop.org/>.
- [25] Digital Video Broadcasting, "ETSI TS 102 823 V1.2.1. Specification for the carriage of synchronised auxiliary data in DVB transport streams." 2014.
- [26] ISO/IEC, "ISO/IEC 13818-1:2013/PDAM 6 Delivery of Timeline for External Data." 2013.
- [27] M. Handley, V. Jacobson, and C. Perkins, "RFC 4566 - SDP: Session Description Protocol," *Request for Comments*, 2006. [Online]. Available: <https://tools.ietf.org/html/rfc4566>.
- [28] B. A. Miller, T. Nixon, C. Tai, and M. D. Wood, "Home networking with Universal Plug and Play," *IEEE Commun. Mag.*, vol. 39, no. 12, pp. 104–109, 2001, doi: 10.1109/35.968819.
- [29] DLNA Consortium, "DLNA Networked Device Interoperability Guidelines," 2009.
- [30] Netflix, "DIAL. Discovery and Launch protocol especificacion. Version 1.7.2." 2015.
- [31] A. Donoho, J. Costa-requena, T. Mcgee, A. Messer, A. Fiddian-green, and J. Fuller, "UPnP™ Device Architecture 1.1," *Architecture*, 2008.
- [32] I. Fette and A. Melnikov, "The WebSocket Protocol. RFC 6455." 2011.
- [33] "Digital Video Broadcasting (DVB); Companion Screens and Streams; Part 2: Content Identification and Media Synchronization. ETSI Standard TS 106 286-2," *Digit. Video Broadcast.*, 2015.

- [34] M. O. Van Deventer, H. Stokking, M. Hammond, J. Le Feuvre, and P. Cesar, "Standards for multi-stream and multi-device media synchronization," *IEEE Commun. Mag.*, vol. 54, no. 3, pp. 16–21, 2016, doi: 10.1109/MCOM.2016.7432166.
- [35] Y. Lim, K. Park, J. Y. Lee, S. Aoki, and G. Fernando, "MMT: An emerging MPEG standard for multimedia delivery over the internet," *IEEE Multimed.*, vol. 20, no. 1, pp. 80–85, 2013, doi: 10.1109/MMUL.2013.7.
- [36] International Organization for Standardization, "ISO/IEC 23008-2:2015. High efficiency coding and media delivery in heterogeneous environment - Part 2: High efficiency video coding." 2015.
- [37] R. Chernock, D. Gómez-Barquero, J. Whitaker, S. I. Park, and Y. Wu, "ATSC 3.0 Next Generation Digital TV Standard - An Overview and Preview of the Issue," *IEEE Transactions on Broadcasting*, vol. 62, no. 1. Institute of Electrical and Electronics Engineers Inc., pp. 154–158, 01-Mar-2016, doi: 10.1109/TBC.2016.2515542.
- [38] F. Boronat, M. Montagud, and V. Vidal, "Smooth control of adaptive media playout to acquire IDMS in cluster-based applications," in *Proceedings - Conference on Local Computer Networks, LCN*, 2011, pp. 613–621, doi: 10.1109/LCN.2011.6115525.
- [39] M. Montagud, F. Boronat, B. Roig, and A. Sapena, "How to perform AMP? Cubic adjustments for improving the QoE," *Comput. Commun.*, vol. 103, pp. 61–73, 2017, doi: 10.1016/j.comcom.2017.01.017.
- [40] B. Rainer and C. Timmerer, "Adaptive media playout for inter-destination media synchronization," in *2013 5th International Workshop on Quality of Multimedia Experience, QoMEX 2013 - Proceedings*, 2013, pp. 44–45, doi: 10.1109/QoMEX.2013.6603206.
- [41] J. Park, P. A. Chou, and J.-N. N. Hwang, "Volumetric Media Streaming for Augmented Reality," in *2018 IEEE Global Communications Conference, GLOBECOM 2018 - Proceedings*, 2018, pp. 1–6, doi: 10.1109/GLOCOM.2018.8647537.
- [42] YouGov, "Global sales of wearables expected to increase by 27% in 2020 | YouGov," 2019. [Online]. Available: <https://yougov.co.uk/topics/resources/articles-reports/2019/11/01/global-sales-wearables-expected-increase-27-2020>. [Accessed: 07-Feb-2021].
- [43] Orion Market Reports, "Head Mounted Display Market 2021: Global Industry Analysis," 2021. [Online]. Available: <https://www.openpr.com/news/2232211/head-mounted-display-market-2021-global-industry-analysis>. [Accessed: 07-Feb-2021].
- [44] S. Rothe, A. Schmidt, M. Montagud, D. Buschek, and H. Hußmann, "Social viewing in cinematic virtual reality: a design space for social movie applications," *Virtual Real.*, vol. 1, p. 3, Oct. 2020, doi: 10.1007/s10055-020-00472-4.
- [45] M. Domanski, O. Stankiewicz, K. Wegner, and T. Grajek, "Immersive visual media — MPEG-I: 360 video, virtual navigation and beyond," in *2017 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2017, pp. 1–9, doi: 10.1109/IWSSIP.2017.7965623.
- [46] J. Jeong, D. Jang, J.-W. Son, and E.-S. Ryu, "Bitrate Efficient 3DoF+ 360 Video View Synthesis for Immersive VR Video Streaming," in *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, 2018, pp. 581–586, doi: 10.1109/ICTC.2018.8539414.

- [47] L. Zelnik-Manor, G. Peters, and P. Perona, "Squaring the circle in panoramas," in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, 2005, pp. 1292-1299 Vol. 2, doi: 10.1109/ICCV.2005.231.
- [48] Z. Chen, Y. Li, and Y. Zhang, "Recent advances in omnidirectional video coding for virtual reality: Projection and evaluation," *Signal Processing*, vol. 146, pp. 66–78, May 2018, doi: 10.1016/J.SIGPRO.2018.01.004.
- [49] E. Kuzyakov and D. Pio, "Next-generation video encoding techniques for 360 video and VR," 2016. [Online]. Available: <https://code.fb.com/virtual-reality/next-generation-video-encoding-techniques-for-360-video-and-vr/>. [Accessed: 27-May-2019].
- [50] J. Le Feuvre and C. Concolato, "Tiled-based adaptive streaming using MPEG-DASH," in *Proceedings of the 7th International Conference on Multimedia Systems - MMSys '16*, 2016, pp. 1–3, doi: 10.1145/2910017.2910641.
- [51] O. A. Niamut, E. Thomas, L. D'Acunto, C. Concolato, F. Denoual, and S. Y. Lim, "MPEG DASH SRD," in *Proceedings of the 7th International Conference on Multimedia Systems - MMSys '16*, 2016, pp. 1–8, doi: 10.1145/2910017.2910606.
- [52] L. Jalal and M. Murrioni, "Enhancing TV broadcasting services: A survey on mulsemmedia quality of experience," in *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2017, pp. 1–7, doi: 10.1109/BMSB.2017.7986192.
- [53] E. B. Saleme, C. A. S. Santos, R. A. Falbo, G. Ghinea, and F. Andres, "Towards a reference ontology on mulsemmedia systems," in *MEDES 2018 - 10th International Conference on Management of Digital EcoSystems*, 2018, pp. 23–30, doi: 10.1145/3281375.3281378.
- [54] E. B. Saleme, C. A. S. Santos, R. A. Falbo, G. Ghinea, and F. Andres, "Mulseonto: A reference ontology to support the design of mulsemmedia systems," *J. Univers. Comput. Sci.*, vol. 25, no. 13, pp. 1761–1786, 2019.
- [55] MPEG, "ISO/IEC 23005. MPEG-V. Media Context and Control." 2010.
- [56] L. Ehley, B. Furht, and M. Ilyas, "Evaluation of multimedia synchronization techniques," in *Proceedings of the International Conference on Multimedia Computing and Systems*, 1994, pp. 514–519, doi: 10.1109/mmcs.1994.292498.
- [57] K. Matsumura, M. J. Evans, Y. Shishikui, and A. McParland, "Personalization of broadcast programs using synchronized internet content," in *ICCE 2010 - 2010 Digest of Technical Papers International Conference on Consumer Electronics*, 2010, pp. 145–146, doi: 10.1109/ICCE.2010.5418823.
- [58] M. Armstrong, J. Barrett, and M. Evans, "WHP185 - Enabling and Enriching Broadcast Services by Combining IP and Broadcast Delivery," no. September, 2010.
- [59] S. Aoki, K. Aoki, H. Hamada, Y. Kanatsugu, M. Yamamoto, and K. Aizawa, "A new transport scheme for hybrid delivery of content over broadcast and broadband," in *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, BMSB 2011 - Conference Programme*, 2011, doi: 10.1109/BMSB.2011.5954888.
- [60] L. B. Yuste, S. M. Saleh Al-Majeed, H. Melvin, and M. Fleury, "Effective synchronisation of Hybrid Broadcast and Broadband TV," in *Digest of Technical Papers - IEEE International Conference on Consumer Electronics*, 2012, pp. 160–161, doi: 10.1109/ICCE.2012.6161788.
- [61] L. Beloqui Yuste and H. Melvin, "Interactive multi-source media synchronization for

- HbbTV," in *MediaSync Workshop*, 2012, pp. 1–8.
- [62] C. Concolato, S. Thomas, R. Bouqueau, and J. Le Feuvre, "Synchronized delivery of multimedia content over uncoordinated broadcast broadband networks," in *MMSys'12 - Proceedings of the 3rd Multimedia Systems Conference*, 2012, pp. 227–232, doi: 10.1145/2155555.2155590.
- [63] GPAC, "GPAC Framework." [Online]. Available: <https://gpac.wp.imt.fr/home/>.
- [64] C. Köhnen, C. Köbel, and N. Hellhund, "A DVB/IP streaming testbed for hybrid digital media content synchronization," in *IEEE International Conference on Consumer Electronics - Berlin, ICCE-Berlin*, 2012, pp. 136–140, doi: 10.1109/ICCE-Berlin.2012.6336493.
- [65] A. Veenhuizen and R. Van Brandenburg, "Frame accurate media synchronization of heterogeneous media sources in an HBB context," *Media Synchronization Work. 2012*, 2012.
- [66] J. Le Feuvre and C. Concolato, "Hybrid Broadcast Services using MPEG DASH," 2013.
- [67] J. Le Feuvre, N. Viet-Thun-Trung, W. Hamidouche, M. Patrick, and D. Pascal, "A Test Bed for Hybrid Broadcast Broadband Services," 2015, p. 4.
- [68] M. Zorrilla, A. Martin, I. Tamayo, S. O'Halpin, and D. Hazael-Massieux, "Reaching devices around an HbbTV television," in *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, BMSB*, 2014, doi: 10.1109/BMSB.2014.6873499.
- [69] C. Howson, E. Gautier, P. Gilberton, A. Laurent, and Y. Legallais, "Second screen TV synchronization," in *Digest of Technical Papers - IEEE International Conference on Consumer Electronics*, 2011, pp. 361–365, doi: 10.1109/ICCE-Berlin.2011.6031815.
- [70] C. Ziegler, "Second screen for HbbTV-Automatic application launch and app-to-app communication enabling novel TV programme related second-screen scenarios," in *Proceedings 2013 IEEE 3rd International Conference on Consumer Electronics - Berlin, ICCE-Berlin 2013*, 2013, doi: 10.1109/ICCE-Berlin.2013.6697990.
- [71] Digital Video Broadcasting, "ETSI TS 106 286-1. Companion Screens and Streams; Part 2: Content Identification and Media Synchronization." .
- [72] V. Vinayagamoorthy, R. Ramdhany, and M. Hammond, "Enabling frame-accurate synchronised companion screen experiences," *TVX 2016 - Proc. ACM Int. Conf. Interact. Exp. TV Online Video*, pp. 83–92, 2016, doi: 10.1145/2932206.2932214.
- [73] C. Ziegler, C. Keimel, R. Ramdhany, and V. Vinayagamoorthy, "On time or not on time: A user study on delays in a synchronised companion-screen experience," in *TVX 2017 - Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video*, 2017, pp. 105–114, doi: 10.1145/3077548.3077557.
- [74] A. Dominguez, M. Agirre, J. Florez, A. Lafuente, I. Tamayo, and M. Zorrilla, "Deployment of a Hybrid Broadcast-Internet Multi-Device Service for a Live TV Programme," *IEEE Trans. Broadcast.*, vol. 64, no. 1, pp. 153–163, Mar. 2018, doi: 10.1109/TBC.2017.2755403.
- [75] M. O. Van Deventer, H. M. Stokking, O. A. Niamut, F. A. Walraven, and V. B. Klos, "Advanced interactive television services require content synchronization," in *Proceedings of IWSSIP 2008 - 15th International Conference on Systems, Signals and Image Processing*, 2008, pp. 109–112, doi: 10.1109/IWSSIP.2008.4604379.

- [76] F. Boronat, J. Lloret, and M. García, "Multimedia group and inter-stream synchronization techniques: A comparative study," *Inf. Syst.*, vol. 34, no. 1, pp. 108–131, Mar. 2009, doi: 10.1016/j.is.2008.05.001.
- [77] M. Montagud, F. Boronat, H. Stokking, and P. Cesar, "Design, development and assessment of control schemes for IDMS in a standardized RTCP-based solution," *Comput. Networks*, vol. 70, pp. 240–259, Sep. 2014, doi: 10.1016/j.comnet.2014.06.004.
- [78] Y. Ishibashi, A. Tsuji, and S. Tasaka, "A group synchronization mechanism for stored media in multicast communications," in *Proceedings of INFOCOM '97*, 1997, vol. 2, pp. 692–700, doi: 10.1109/infcom.1997.644522.
- [79] Y. Ishibashi and S. Tasaka, "Group synchronization mechanism for live media in multicast communications," in *Conference Record / IEEE Global Telecommunications Conference*, 1997, vol. 2, pp. 746–752, doi: 10.1109/glocom.1997.638431.
- [80] R. Arima, M. Sithu, and Y. Ishibashi, "QoE Assessment of Fairness between Players in Networked Virtual 3D Objects Identification Game Using Haptic, Olfactory, and Auditory Senses," *Int. J. Commun. Netw. Syst. Sci.*, vol. 10, no. 07, pp. 129–141, 2017, doi: 10.4236/ijcns.2017.107007.
- [81] M. Mauve, J. Vogel, V. Hilt, and W. Effelsberg, "Local-lag and timewarp: Providing consistency for replicated continuous applications," *IEEE Trans. Multimed.*, vol. 6, no. 1, pp. 47–57, 2004, doi: 10.1109/TMM.2003.819751.
- [82] I. Vaishnavi, P. Cesar, D. Bulterman, O. Friedrich, S. Gunkel, and D. Geerts, "From IPTV to synchronous shared experiences challenges in design: Distributed media synchronization," *Signal Process. Image Commun.*, vol. 26, no. 7, pp. 370–377, 2011, doi: 10.1016/j.image.2011.01.006.
- [83] F. Boronat, J. C. G. Cebollada, and J. L. Mauri, "An RTP/RTCP based approach for multimedia group and inter-stream synchronization," *Multimed. Tools Appl.*, vol. 40, no. 2, pp. 285–319, 2008, doi: 10.1007/s11042-008-0208-1.
- [84] R. van Brandenburg, H. Stokking, F. Boronat, M. Mario Montagud, and K. Gross, "Inter-destination Media Synchronization using the RTP Control Protocol(RTCP)," no. draft-ietf-avtcore-idms-06, 2012.
- [85] B. Rainer and C. Timmerer, "Self-organized Inter-Destination Multimedia Synchronization for adaptive media streaming," in *MM 2014 - Proceedings of the 2014 ACM Conference on Multimedia*, 2014, pp. 327–336, doi: 10.1145/2647868.2654938.
- [86] F. Boronat, M. Montagud, P. Salvador, and J. Pastor, "Wersync: A web platform for synchronized social viewing enabling interaction and collaboration," *J. Netw. Comput. Appl.*, vol. 175, p. 102939, Feb. 2021, doi: 10.1016/j.jnca.2020.102939.
- [87] M. Montagud and F. Boronat, "Enhanced adaptive RTCP-based Inter-Destination Multimedia Synchronization approach for distributed applications," *Comput. Networks*, vol. 56, no. 12, pp. 2912–2933, 2012, doi: 10.1016/j.comnet.2012.05.003.
- [88] N. Ducheneaut, R. J. Moore, L. Oehlberg, J. D. Thornton, and E. Nickell, "Social TV: Designing for distributed, sociable television viewing," *Int. J. Hum. Comput. Interact.*, vol. 24, no. 2, pp. 136–154, 2008, doi: 10.1080/10447310701821426.
- [89] J. D. Weisz, S. Kiesler, H. Zhang, Y. Ren, R. E. Kraut, and J. A. Konstan, "Watching together:

- Integrating text chat with video," in *Conference on Human Factors in Computing Systems - Proceedings*, 2007, pp. 877–886, doi: 10.1145/1240624.1240756.
- [90] D. A. Shamma, M. Bastéa-Forte, N. Joubert, and Y. Liu, "Enhancing online personal connections through the synchronized sharing of online video," in *Conference on Human Factors in Computing Systems - Proceedings*, 2008, pp. 2931–2936, doi: 10.1145/1358628.1358786.
- [91] D. Geerts, I. Vaishnavi, R. Mekuria, O. Van Deventer, and P. Cesar, "Are we in sync? Synchronization requirements for watching online video together," in *Conference on Human Factors in Computing Systems - Proceedings*, 2011, pp. 311–314, doi: 10.1145/1978942.1978986.
- [92] D. Geerts, "Comparing voice chat and text chat in a communication tool for interactive television," in *ACM International Conference Proceeding Series*, 2006, vol. 189, pp. 461–464, doi: 10.1145/1182475.1182537.
- [93] E. M. Huang *et al.*, "Of social television comes home: A field study of communication choices and practices in TV-based text and voice chat," in *Conference on Human Factors in Computing Systems - Proceedings*, 2009, pp. 585–594, doi: 10.1145/1518701.1518792.
- [94] R. Mekuria, P. Cesar, and D. Bulterman, "Digital TV: The effect of delay when watching football," in *EuroiTV'12 - Proceedings of the 10th European Conference on Interactive TV and Video*, 2012, pp. 71–74, doi: 10.1145/2325616.2325632.
- [95] B. Rainer, S. Petschnig, C. Timmerer, and H. Hellwagner, "Is one second enough? Evaluating QoE for inter-destination multimedia synchronization using human computation and crowdsourcing," in *2015 7th International Workshop on Quality of Multimedia Experience, QoMEX 2015*, 2015, doi: 10.1109/QoMEX.2015.7148107.
- [96] M. Hosseini and V. Swaminathan, "Adaptive 360 VR Video Streaming: Divide and Conquer," in *2016 IEEE International Symposium on Multimedia (ISM)*, 2016, pp. 107–110, doi: 10.1109/ISM.2016.0028.
- [97] M. Graf, C. Timmerer, and C. Mueller, "Towards Bandwidth Efficient Adaptive Streaming of Omnidirectional Video over HTTP," in *Proceedings of the 8th ACM on Multimedia Systems Conference - MMSys'17*, 2017, pp. 261–271, doi: 10.1145/3083187.3084016.
- [98] S. Petrangeli, V. Swaminathan, M. Hosseini, and F. De Turck, "An HTTP/2-Based Adaptive Streaming Framework for 360° Virtual Reality Videos," in *Proceedings of the 2017 ACM on Multimedia Conference - MM '17*, 2017, pp. 306–314, doi: 10.1145/3123266.3123453.
- [99] X. Corbillon, F. De Simone, G. Simon, and P. Frossard, "Dynamic adaptive streaming for multi-viewpoint omnidirectional videos," in *Proceedings of the 9th ACM Multimedia Systems Conference on - MMSys '18*, 2018, pp. 237–249, doi: 10.1145/3204949.3204968.
- [100] D. Ochi, Y. Kunita, A. Kameda, A. Kojima, and S. Iwaki, "Live streaming system for omnidirectional video," in *2015 IEEE Virtual Reality (VR)*, 2015, pp. 349–350, doi: 10.1109/VR.2015.7223439.
- [101] C. Velasco and M. Obrist, "Multisensory Experiences: Where the senses meet technology," *Oxford*, p. 112, 2020, doi: 10.1093/oso/9780198849629.001.0001.
- [102] H. Q. Dinh, N. Walker, L. F. Hodges, Chang Song, and A. Kobayashi, "Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual

- environments," in *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*, 2003, pp. 222–228, doi: 10.1109/vr.1999.756955.
- [103] T. Bi, R. Lyons, G. Fox, and G.-M. Muntean, "Improving Student Learning Satisfaction by Using an Innovative DASH-based Multiple Sensorial Media Delivery Solution," *IEEE Trans. Multimed.*, pp. 1–1, 2020, doi: 10.1109/tmm.2020.3025669.
- [104] A. A. Alkasasbeh and G. Ghinea, "Using olfactory media cues in e-learning – perspectives from an empirical investigation," *Multimed. Tools Appl.*, vol. 79, no. 27–28, pp. 19265–19287, 2020, doi: 10.1007/s11042-020-08763-3.
- [105] Y. Sulema, "Mulsemmedia vs. Multimedia: State of the art and future trends," in *International Conference on Systems, Signals, and Image Processing*, 2016, vol. 2016-June, pp. 1–5, doi: 10.1109/IWSSIP.2016.7502696.
- [106] N. Murray, O. A. Ademoye, G. Ghinea, and G.-M. Muntean, "A Tutorial for Olfaction-Based Multisensorial Media Application Design and Evaluation," *ACM Comput. Surv.*, vol. 50, no. 5, pp. 1–30, Sep. 2017, doi: 10.1145/3108243.
- [107] E. B. Saleme, A. Covaci, G. Mesfin, C. A. S. Santos, and G. Ghinea, "Mulsemmedia DIY: A survey of devices and a tutorial for building your own mulsemmedia environment," *ACM Computing Surveys*, vol. 52, no. 3. Association for Computing Machinery, 01-Jun-2019, doi: 10.1145/3319853.
- [108] A. Covaci *et al.*, "360° Mulsemmedia: A way to improve subjective QoE in 360° videos," *MM 2019 - Proc. 27th ACM Int. Conf. Multimed.*, pp. 2378–2386, 2019, doi: 10.1145/3343031.3350954.
- [109] D. Hajas, D. Ablart, O. Schneider, and M. Obrist, "I can feel it moving: Science Communicators Talking About the Potential of Mid-Air Haptics," *Front. Comput. Sci.*, vol. 2, p. 48, 2020, doi: 10.3389/fcomp.2020.534974.
- [110] C. Velasco *et al.*, "Multisensory Approaches to Human-Food Interaction," in *ICMI 2020 - Proceedings of the 2020 International Conference on Multimodal Interaction*, 2020, vol. 3, no. 20, pp. 878–880, doi: 10.1145/3382507.3419749.
- [111] C. T. Vi *et al.*, "LeviSense: A platform for the multisensory integration in levitating food and insights into its effect on flavour perception," *Int. J. Hum. Comput. Stud.*, vol. 139, p. 102428, Jul. 2020, doi: 10.1016/j.ijhcs.2020.102428.
- [112] Z. Akhtar, K. Siddique, A. Rattani, S. L. Lutfi, and T. H. Falk, "Why is multimedia quality of experience assessment a challenging problem?," *IEEE Access*, vol. 7, pp. 117897–117915, 2019, doi: 10.1109/ACCESS.2019.2936470.
- [113] M. Melo *et al.*, "Immersive multisensory virtual reality technologies for virtual tourism: A study of the user's sense of presence, satisfaction, emotions, and attitudes," *Multimed. Syst.*, 2022, doi: 10.1007/s00530-022-00898-7.
- [114] International Organization for Standardization, "ISO 5496:2006 Sensory analysis — Methodology — Initiation and training of assessors in the detection and recognition of odours." 2006.
- [115] O. A. Ademoye, N. Murray, G.-M. Muntean, and G. Ghinea, "Audio Masking Effect on Inter-Component Skews in Olfaction-Enhanced Multimedia Presentations," *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 12, no. 4, pp. 1–14, Aug. 2016, doi:

- 10.1145/2957753.
- [116] G. Mesfin, N. Hussain, E. Kani-Zabihi, A. Covaci, E. B. Saleme, and G. Ghinea, "QoE of cross-modally mapped Mulsemedia: an assessment using eye gaze and heart rate," *Multimed. Tools Appl.*, vol. 79, no. 11–12, pp. 7987–8009, 2020, doi: 10.1007/s11042-019-08473-5.
- [117] E. B. Saleme *et al.*, "The influence of human factors on 360° mulsemedia QoE," *Int. J. Hum. Comput. Stud.*, vol. 146, p. 102550, Feb. 2021, doi: 10.1016/j.ijhcs.2020.102550.
- [118] B. S. Choi, S. H. Joo, and H. Y. Lee, "Sensory effect metadata for SMMD media service," in *Proceedings of the 2009 4th International Conference on Internet and Web Applications and Services, ICIW 2009*, 2009, pp. 649–654, doi: 10.1109/ICIW.2009.104.
- [119] J. K. Yun, J. H. Jang, and K. D. Moon, "Five sense media playback technology using multile devices synchronization," in *Proceeding - 5th International Conference on Computer Sciences and Convergence Information Technology, ICCIT 2010*, 2010, pp. 73–75, doi: 10.1109/ICCIT.2010.5711032.
- [120] K. Hirota, S. Ebisawa, T. Amemiya, and Y. Ikei, "A theater for viewing and editing multi-sensory content," in *ISVRI 2011 - IEEE International Symposium on Virtual Reality Innovations 2011, Proceedings*, 2011, pp. 239–244, doi: 10.1109/ISVRI.2011.5759643.
- [121] M. Walzl, B. Rainer, C. Timmerer, and H. Hellwagner, "A toolset for the authoring, simulation, and rendering of sensory experiences," in *MM 2012 - Proceedings of the 20th ACM International Conference on Multimedia*, 2012, pp. 1469–1472, doi: 10.1145/2393347.2396522.
- [122] M. Walzl, B. Rainer, and H. Hellwagner, "An end-to-end tool chain for Sensory Experience based on MPEG-V," *Signal Process. Image Commun.*, vol. 28, no. 2, pp. 136–150, Feb. 2013, doi: 10.1016/J.IMAGE.2012.10.009.
- [123] Z. Yuan, G. Ghinea, and G. M. Muntean, "Quality of experience study for multiple sensorial media delivery," in *IWCMC 2014 - 10th International Wireless Communications and Mobile Computing Conference*, 2014, pp. 1142–1146, doi: 10.1109/IWCMC.2014.6906515.
- [124] M. Walzl, C. Timmerer, B. Rainer, and H. Hellwagner, "Sensory effects for ambient experiences in the World Wide Web," *Multimed. Tools Appl.*, vol. 70, no. 2, pp. 1141–1160, May 2014, doi: 10.1007/s11042-012-1099-8.
- [125] E. B. Saleme and C. A. S. Santos, "PlaySEM: A platform for rendering MulSeMedia compatible with MPEG-V," in *WebMedia 2015 - Proceedings of the 21st Brazilian Symposium on Multimedia and the Web*, 2015, pp. 145–148, doi: 10.1145/2820426.2820450.
- [126] Z. Yuan, G. Ghinea, and G. M. Muntean, "Beyond multimedia adaptation: Quality of experience-aware multi-sensorial media delivery," *IEEE Trans. Multimed.*, vol. 17, no. 1, pp. 104–117, Jan. 2015, doi: 10.1109/TMM.2014.2371240.
- [127] S. H. Shin, K. S. Ha, H. O. Yun, and Y. S. Nam, "Realistic media authoring tool based on MPEG-V international standard," in *International Conference on Ubiquitous and Future Networks, ICUFN*, 2016, vol. 2016-Augus, pp. 730–732, doi: 10.1109/ICUFN.2016.7537133.
- [128] E. B. Saleme, C. A. S. Santos, and G. Ghinea, "Coping with the Challenges of Delivering Multiple Sensorial Media," *IEEE Multimedia*, pp. 1–1, 2018.
- [129] A. Covaci, L. Zou, I. Tal, G. M. Muntean, and G. Ghinea, "Is multimedia multisensorial? - A

- review of mulsemmedia systems," *ACM Computing Surveys*, vol. 51, no. 5. Association for Computing Machinery, 01-Aug-2018, doi: 10.1145/3233774.
- [130] E. B. Saleme, C. A. S. Santos, and G. Ghinea, "A mulsemmedia framework for delivering sensory effects to heterogeneous systems," *Multimed. Syst.*, p. 27, May 2019, doi: 10.1007/s00530-019-00618-8.
- [131] I. S. Comsa *et al.*, "Do i Smell Coffee? The Tale of a 360° Mulsemmedia Experience," *IEEE Multimed.*, vol. 27, no. 1, pp. 27–36, 2020, doi: 10.1109/MMUL.2019.2954405.
- [132] MPEG, "MPEG-7 Part 5: MDS, ISO/IEC 15938-5." 2003.
- [133] M. Waltl, C. Timmerer, and H. Hellwagner, "Improving the quality of multimedia experience through sensory effects," in *2010 2nd International Workshop on Quality of Multimedia Experience, QoMEX 2010 - Proceedings*, 2010, pp. 124–129, doi: 10.1109/QOMEX.2010.5517704.
- [134] M. Waltl, C. Timmerer, and H. Hellwagner, "A test-bed for quality of multimedia experience evaluation of sensory effects," in *2009 International Workshop on Quality of Multimedia Experience, QoMEX 2009*, 2009, pp. 145–150, doi: 10.1109/QOMEX.2009.5246962.
- [135] N. Ranasinghe, K.-Y. Lee, G. Suthokumar, and E. Y.-L. Do, "Virtual ingredients for food and beverages to create immersive taste experiences," *Multimed. Tools Appl.*, vol. 75, no. 20, pp. 12291–12309, Oct. 2016, doi: 10.1007/s11042-015-3162-8.
- [136] J. K. Yun, M. G. Kim, J. H. Jang, and K. R. Park, "Media/playback device synchronization for the 4D broadcasting service system," in *Digest of Technical Papers - IEEE International Conference on Consumer Electronics*, 2012, pp. 329–330, doi: 10.1109/ICCE.2012.6161891.
- [137] K. Yoon, B. Choi, E. S. Lee, and T. B. Lim, "4-D broadcasting with MPEG-V," in *2010 IEEE International Workshop on Multimedia Signal Processing, MMSP2010*, 2010, pp. 257–262, doi: 10.1109/MMSP.2010.5662029.
- [138] J. K. Yun, J. H. Jang, and K. D. Moon, "Development of the real-sense media broadcasting service system based on the SMMD," in *Digest of Technical Papers - IEEE International Conference on Consumer Electronics*, 2011, pp. 435–436, doi: 10.1109/ICCE.2011.5722669.
- [139] L. Jalal, M. Anedda, V. Popescu, and M. Murrioni, "Internet of Things for Enabling Multi Sensorial TV in Smart Home," in *2018 IEEE Broadcast Symposium, BTS 2018*, 2018, pp. 1–5, doi: 10.1109/BTS.2018.8550959.
- [140] L. Jalal, M. Anedda, V. Popescu, and M. Murrioni, "QoE Assessment for Broadcasting Multi Sensorial Media in Smart Home Scenario," in *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, BMSB*, 2018, vol. 2018-June, pp. 1–5, doi: 10.1109/BMSB.2018.8436875.
- [141] J. K. Yun, J. H. Jang, K. R. Park, and D. W. Han, "Real-sense media representation technology using multiple devices synchronization," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2009, vol. 5860 LNCS, pp. 343–353, doi: 10.1007/978-3-642-10265-3_31.
- [142] N. Murray, B. Lee, Y. Qiao, and G.-M. Muntean, "Multiple-Scent Enhanced Multimedia Synchronization," *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 11, no. 1s, pp. 1–28, Oct. 2014, doi: 10.1145/2637293.
- [143] N. Murray, G. M. Muntean, Y. Qiao, and B. Lee, "Olfaction-enhanced multimedia

- synchronization," in *MediaSync: Handbook on Multimedia Synchronization*, Cham: Springer International Publishing, 2018, pp. 319–356.
- [144] F. Boronat, M. Montagud, H. Stokking, and O. Niamut, "The need for inter-destination synchronization for emerging social interactive multimedia applications," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 150–158, 2012, doi: 10.1109/MCOM.2012.6353695.
- [145] J. Belda, Montagud M., F. Boronat, M. Martinez, and J. Pastor, "Wersync: A WEB-Based Platform for Distributed Media Synchronization and Social Interaction," *Proc. ACM Int. Conf. Interact. Exp. Telev. Online Video 2015*, pp. 9–10, 2015.
- [146] NIELSEN, "The Nielsen Comparable Metrics Report: Q4 2016," 2017.
- [147] W. J. Kooij, H. M. Stokking, R. Van Brandenburg, and P. T. De Boer, "Playout delay of tv signals: Measurement system design, validation and results," in *TVX 2014 - Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and Online Video*, 2014, pp. 23–30, doi: 10.1145/2602299.2602310.
- [148] R. N. Mekuria, "Inter-destination media synchronization for TV broadcasts:: TU Delft Institutional Repository," 2011. [Online]. Available: <https://repository.tudelft.nl/islandora/object/uuid:61907582-3fe2-4e1f-86af-47c0eacfb04>.
- [149] R. Braden, D. Borman, and C. Partridge, "Computing the internet checksum," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 19, no. 2, pp. 86–94, Apr. 1989, doi: 10.1145/378444.378453.
- [150] I. Grigorik, *High-Performance Browser Networking*. 2013.
- [151] Z. Yuan, T. Bi, G. M. Muntean, and G. Ghinea, "Perceived synchronization of mulsemmedia services," *IEEE Trans. Multimed.*, vol. 17, no. 7, pp. 957–966, Jul. 2015, doi: 10.1109/TMM.2015.2431915.
- [152] FFmpeg, "FFmpeg." [Online]. Available: <https://ffmpeg.org/>. [Accessed: 22-Mar-2019].
- [153] International Telecommunication Union (ITU), "Recommendation ITU-R BT.470, Conventional Television Systems," 1998.
- [154] R. Pantos and W. May, "HTTP Live Streaming," *IETF Draft*, pp. 1–33, 2012.
- [155] H. Schulzrinne, A. Rao, and R. Lanphier, "Real Time Streaming Protocol [RFC 2326]," *Internet Eng. Task Force*, vol. 2326, pp. 1–283, 1998.
- [156] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RFC 3550 - Real-time Transport Protocol (RTP)," *IETF: Network Working Group*. 2003.
- [157] D. Gómez, F. Boronat, M. Montagud, and C. Luzón, "End-to-end DASH platform including a network-based and client-based adaptive quality switching module," in *Proceedings of the 7th International Conference on Multimedia Systems, MMSys 2016*, 2016, pp. 317–320, doi: 10.1145/2910017.2910638.
- [158] M. Zorrilla, N. Borch, F. Daoust, A. Erk, J. Flórez, and A. Lafuente, "A Web-based distributed architecture for multi-device adaptation in media applications," *Pers. Ubiquitous Comput.*, vol. 19, no. 5–6, pp. 803–820, 2015, doi: 10.1007/s00779-015-0864-x.
- [159] A. Dominguez, J. Florez, A. Lafuente, S. Masneri, I. Tamayo, and M. Zorrilla, "A methodology

- for User Interface adaptation of multi-device broadcast-broadband services," *IEEE Access*, 2020, doi: 10.1109/ACCESS.2020.3039616.
- [160] M. Montagud, F. Boronat, J. Belda, and D. Marfil, "Use of web components to develop interactive, customizable and multi-device video consumption platforms," in *Communications in Computer and Information Science*, 2016, vol. 605, pp. 26–43, doi: 10.1007/978-3-319-38907-3_4.
- [161] DASH Industry Forum, "Dash.js." [Online]. Available: <https://dashif.org/dash.js/>. [Accessed: 24-May-2019].
- [162] three.js, "three.js - Javascript 3D library." [Online]. Available: <https://threejs.org/>. [Accessed: 22-Mar-2019].
- [163] R. Steinmetz, "Human perception of jitter and media Synchronization," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 1, pp. 61–72, 1996, doi: 10.1109/49.481694.
- [164] Can I Use, "Can I Use: HEVC." [Online]. Available: <https://caniuse.com/#search=hevc>.
- [165] Can I Use, "Can I Use: MPEG4." [Online]. Available: <https://caniuse.com/#search=MPEG-4>.
- [166] ITU, "Methodology for the subjective assessment of the quality of television pictures," *ITU-R Bt.500-13*, vol. 211, no. BT.500-13, pp. 1–48, 2012.
- [167] J. Brooke, "SUS-A quick and dirty usability scale," *Usability Eval. Ind.*, vol. 189, no. 194, pp. 4–7, 1996.
- [168] Itu-T P.800, "Methods for objective and subjective assessment of quality," *Int. Telecommun. Union*, 1998.
- [169] J. Brooke, "SUS - A quick and dirty usability scale Industrial usability evaluation," in *Usability evaluation in industry*, I. Jordan, P. W., Thomas, B. A. Weerdmeester and McClelland, Ed. London: Taylor & Francis, 2011, pp. 189–194.
- [170] A. Bangor, P. Kortum, and J. Miller, "Determining what individual SUS scores mean: adding an adjective rating scale," *Usability Stud.*, vol. 4, no. 3, pp. 114–123, 2009.
- [171] D. Marfil, F. Boronat, J. López, and B. Roig, "Mechanism to validate and compare inter-device / destination media synchronization solutions adopted in multi-screen scenarios," *IEEE Lat. Am. Trans.*, vol. 100, no. 1e (Early Access), 2020.

Anexo PUBLICACIONES Y TRABAJOS DERIVADOS DE LA TESIS DOCTORAL

A continuación, se listan los resultados y trabajos derivados de esta Tesis Doctoral.

PUBLICACIONES EN REVISTAS DE INVESTIGACIÓN

- **D. Marfil**, F. Boronat, J. Gonzalez, A. Sapena, "Integration of multi-sensorial effects in synchronized immersive hybrid TV escenarios", IEEE Access (under review).
- **D. Marfil**, F. Boronat, J. González, A. Vidal, "Design and assessment of a scalable and customizable low-cost tiled display system", IEEE Latin Am. Trans. (Accepted in 2020) (in Spanish) , ISSN: 1548-0992, Q4.
<https://latamt.ieee9.org/index.php/transactions/article/view/3647/876>
- **D. Marfil**, F. Boronat, J. López, B. Roig, "Mechanism to validate and compare inter-device/destination media synchronization solutions adopted in multi-screen scenarios", IEEE Latin Am. Trans. (Accepted in 2020) (in Spanish), ISSN: 1548-0992, Q4.
<https://latamt.ieee9.org/index.php/transactions/article/view/3374/558>
- Montagud, M., Boronat, F., Pastor, J., **D. Marfil**. Web-based platform for a customizable and synchronized presentation of subtitles in single- and multi-screen scenarios. *Multimed Tools Appl* (2020). ISSN 1380-7501, <https://doi.org/10.1007/s11042-020-08955-x>
- **D. Marfil**, F. Boronat, J. López and A. Vidal, "Enhancing the Broadcasted TV Consumption Experience With Broadband Omnidirectional Video Content," in *IEEE Access*, vol. 7, pp. 171864-171883, 2019.
DOI: 10.1109/ACCESS.2019.2956084
- **D. Marfil**; Boronat, Fernando; Sapena Piera, Almanzor; Vidal Meló, Anna. (2019) "Synchronization Mechanisms for Multi-User and Multi-Device Hybrid Broadcast and Broadband Distributed Scenarios". IEEE Access (7)605 - 624. DOI: 10.1109/ACCESS.2018.2885580
- **D. Marfil**; Boronat, Fernando; Mario Montagud; Sapena Piera, Almanzor. (2019) "IDMS solution for hybrid broadcast broadband delivery within the context of HbbTV standard". IEEE Trans. on Broadcasting, vol. 65, no. 4, pp. 645-663, 2019. DOI: 10.1109/TBC.2018.287828
- F. Boronat, **D. Marfil**, M. Montagud, F. Pastor, "HbbTV-compliant Platform for Hybrid Media Delivery and Synchronization on Single- and Multi-Device Scenarios", IEEE Trans. on Broadcasting, vol. 64, no. 3, pp. 721-746, 2018. DOI: 10.1109/TBC.2017.2781124
- F. Boronat, M. Montagud, **D. Marfil**, C. Luzón, "Hybrid Broadcast/Broadband TV Services and Media Synchronization. Demands, preferences and Expectations of Spanish Consumers", IEEE Transactions on Broadcasting, vol. 64, no. 1, pp. 52-69, 2018. DOI: 10.1109/TBC.2017.2737819
- M. Montagud, P. Cesar, F. Boronat, **D. Marfil**, "Social Media Usage Combined with TV/Video Watching: Opportunities and Associated Challenges; Towards Augmented, Interactive, Personalized and Shared Experiences", IEEE Computer Society Special Technical Community on Social Networking (STCSN) E-Letter, vol. 3, no. 2, 2015. Persistent identifier: urn:NBN:nl:ui:18-23726

PUBLICACIONES EN CONGRESOS/WORKSHOPS DE INVESTIGACIÓN

- **D. Marfil**, F. Boronat, J. López, A. Sapena, "Mecanismo para la evaluación de la sincronización en la presentación de contenidos multi-pantalla", XIV Jornadas de Ingeniería Telemática (JITEL 2019), Zaragoza (Spain), 22-24 October 2019.
- J. López, **D. Marfil**, F. Boronat, J. Pastor, "Escenario virtual inmersivo basado en web para el consumo de contenido omnidireccional", XIV Jornadas de Ingeniería Telemática (JITEL 2019), Zaragoza (Spain), 22-24 October 2019.
- **D. Marfil**, F. Boronat, "*Sincronización híbrida entre Destinatarios*", IV Workshop QoS y QoE en Comunicación Multimedia (QQCM18), Zaragoza (Spain), 12-13 Julio 2018.
- **D. Marfil**, F. Boronat, M. Montagud, P. Salvador, "*Plataforma extremo-a-extremo compatible con el estándar HbbTV 2.0 para la TV híbrida y multi-dispositivo*", XIII Jornadas de Ingeniería Telemática (JITEL 2017), Valencia (Spain), 27-29 Septiembre 2017.
- P. Salvador, F. Boronat, M. Montagud, **D. Marfil**, "*Sistema videowall de bajo coste basado en Raspberry Pi, personalizable y configurable dinámica y remotamente vía Web*", XIII Jornadas de Ingeniería Telemática (JITEL 2017), Valencia (Spain), 27-29 Septiembre 2017.
- F. Boronat, **D. Marfil**, M. Montagud, C. Luzón, "*An Hbbtv-Compliant End-To-End Plaform to Enable Hybrid Multi-Device TV Scenarios*", 9th edition of NEM SUMMIT 2016 Conference - Immersive and Interactive Media Experiences Session, Noviembre 2016.
- F. Boronat, M. Montagud, M. Martínez, **D. Marfil**, "*Estudio sobre necesidades, hábitos, preferencias y expectativas de los usuarios con respecto a la sincronización multimedia en escenarios híbridos.*", XII Jornadas de Ingeniería Telemática (JITEL), Octubre 2015, Mallorca.
- F. Boronat, M. Montagud, M. Martínez, **D. Marfil**, "*Preferencias, necesidades y expectativas de los usuarios españoles en escenarios multimedia híbridos broadcast/broadband*", Congreso TV Digital Interactiva – IV Jornadas Iberoamericanas de Aplicaciones y Usabilidad de la TVDi, Octubre 2015, Mallorca.

CAPÍTULOS DE LIBRO

- M. Montagud, F. Boronat, J. Belda, **D. Marfil**, "*Use of Web Components to Develop Interactive, Customizable and Multi-Device Video Consumption Platforms*". Applications and Usability of Interactive TV, Volume 605 of the series Communications in Computer and Information Science (CCIS), Springer-Verlag, pp. 26-43, Abril 2016.