



UNIVERSIDAD
POLITECNICA
DE VALENCIA



Máster Universitario
en Tecnologías, Sistemas y
Redes de Comunicaciones

Codificación y Transmisión de videos 360° en proyección Equirectangular utilizando DASH en Sistemas de Comunicaciones MIMO

Autor: Miguel Fernández Dasí

Director 1: Alberto González Salvador

Director 2: M^a Ángeles Simarro Haro

Fecha de comienzo: 12/07/2021

Lugar de trabajo: Grupo de Tratamiento de Audio y Comunicaciones del iTEAM

Objetivos

En este trabajo se quiere introducir el uso de video 360° en el contexto de las carreras de drones y para ello se propone una plataforma que, mediante el uso de MPEG-DASH, pueda ofrecer este video a los clientes manteniendo la mejor calidad posible. Además, se compara el uso de diferentes configuraciones CMP para reducir el ancho de banda requerido cuando se ofrece el contenido como video bajo demanda.

Metodología

En este trabajo se ha utilizado una metodología de trabajo ‘ágil’, donde se han realizado diferentes iteraciones en cada una de las fases hasta que se ha obtenido un resultado satisfactorio.

Desarrollos teóricos realizados

En este trabajo se ha definido una plataforma extremo-a-extremo con la que transmitir video 360° desde un dron en el contexto de una carrera de drones a los usuarios finales que consumirán el contenido a través de Internet. Dentro de la plataforma se ha definido como se realizarán las diferentes transformaciones del video y como se realizaría una agregación de streams de video para realizar cambios multicámara.

Desarrollo de prototipos y trabajo de laboratorio

Durante el desarrollo de este trabajo se ha creado un entorno de pruebas con una cámara 360° y un ordenador que ha actuado de servidor de contenidos, tanto para ofrecer el contenido via streaming como contenido bajo demanda. Con ello, se han realizado diferentes pruebas para evaluar el rendimiento ofrecido por este banco de pruebas.

Resultados

Con los resultados obtenidos se ha podido comprobar que es posible transmitir video 360° en el contexto de una carrera de drones para que los clientes puedan consumir el contenido via streaming. Además, también se ha visto que se consigue un buen resultado en cuanto a calidad, retardo y throughput al utilizar diferentes configuraciones de la proyección CMP.

Líneas futuras

Como trabajo futuro, se debería trabajar investigando el rendimiento que ofrecen otras configuraciones de la proyección CMP para observar si alguna de ellas podría dar mejores resultados que las vistas en este trabajo. Por otro lado, dado que la salida de video de las cámaras 360° es en proyección ERP, se debería comparar el rendimiento que se obtiene al incluir zonas de interés en la imagen que concentren la calidad en áreas concretas. De esta forma se consigue ofrecer una mejor calidad al cliente en las regiones donde ocurra la acción. Por último, dado que las cámaras de los drones no se colocan paralelas al suelo para compensar la inclinación de este cuando acelera, se debería trabajar en incluir este aspecto a la hora de crear las zonas de interés (en el caso de ERP) o las diferentes caras (en el caso de CMP) para asegurarse que coinciden de forma correcta con lo que se está grabando.

Abstract

Today virtual reality is acquiring an increasingly important role in our society, thanks to improvements in communication systems it is possible to achieve transmission speeds high enough to be able to view content on demand or via streaming without stops and fluid form.

This is why the use of virtual reality helmets in different areas is increasing in popularity, mainly due to the immersion it offers the user, making them feel closer to the content they are viewing. For these reasons, the inclusion of 360° video in the transmission of the image in drone races is a natural step, which improves the experience of the user and the pilot himself.

Research in the search for ways to transmit content more efficiently is an objective that encompasses numerous professional fields and that not only improves the quality of drone racing, but also that of any content that is consumed in virtual reality.

Autor: Fernández Dasí Miguel, email: miferda@teleco.upv.es

Director 1: González Salvador Alberto, email: agonzal@dcom.upv.es

Director 2: Simarro Haro M^a Ángeles, email: mdesiha@iteam.upv.es

Fecha de entrega: 12-07-21

ÍNDICE

I. Introducción	4
II. Estado del arte	7
II.1. Transmisión del video en carreras de drones.....	7
II.2. Head-Mounted Display	8
II.3. Consumo de contenidos en streaming y VoD	9
II.4. Dynamic Adaptative Streaming over HTTP	12
III. Evaluación de Costes del Proyecto	14
III.1. Costes económicos.....	14
III.2. Costes temporales	14
IV. Evaluación de Costes del Proyecto	16
IV.1. Configuración CMP 1:1.....	17
IV.1. Configuración CMP 11:6.....	18
V. Esquema de distribución de contenido	20
VI. Pruebas realizadas	22
VI.1. Prueba de plataforma de streaming.....	24
VI.1.1. Prueba de reproducción de contenido via streaming.....	24
VI.1.2. Prueba de reproducción de contenido bajo demanda.....	26
VI.2. Conclusiones de pruebas realizadas.....	29
VII. Conclusión y líneas de trabajo futuras	30
VII.1.Líneas de trabajo futuras.....	30
Referencias	32

I. INTRODUCCIÓN

Desde la aparición de las primeras plataformas para el consumo de contenidos bajo demanda (VoD) siempre se ha buscado ofrecer la mejor calidad posible para el usuario. Con el avance de las tecnologías de comunicaciones se ha podido aumentar cada vez más la calidad del contenido transmitido, y hoy en día, en plataformas como YouTube o Netflix se ofrecen contenidos con resoluciones de hasta 4K (3840x2160).

En los últimos años, con la aparición de HMDs (Head-Mounted Display) orientados al consumidor, los contenidos 360° comenzaron a popularizarse. En concreto, los contenidos 360° requieren de una mayor resolución para poder ofrecer una calidad similar a la de un contenido en 2 dimensiones. Debido a esto, y al hecho de que se requiere de un HMD para poder visualizar el contenido, estos contenidos se presentan más restrictivos que el video tradicional. Sin embargo, esto no ha impedido que su consumo haya aumentado y que compañías como Oculus, HTC, Valve o HP sigan sacando al mercado nuevos HMD. En la Fig. 1 se pueden observar primeras versiones de HMD de Oculus y HTC.

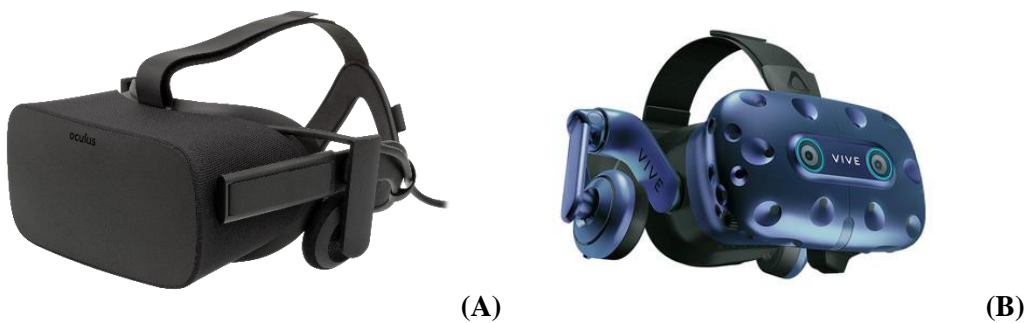


Fig. 1: Primeras iteraciones de HMD de Oculus (A) y HTC (B).

Principalmente, el contenido 360° se asocia con experiencias de realidad virtual (RV), ya sea para jugar a un videojuego o para simular que el usuario se encuentra en un cine, teatro, etc. Son las experiencias de videojuegos las que mejor han aprovechado las posibilidades que ofrece la RV, haciendo que sean hoy en día una parte muy importante de las ventas de aplicaciones de RV.

Sin embargo, debido a la situación propiciada por la pandemia, la RV ha recibido una mayor importancia este último año. La gran mayoría de aplicaciones para realizar reuniones o videoconferencias se han adaptado para ofrecer una versión en RV, sirviéndose de la mejora en interactividad e inmersión de los HMDs para mejorar la experiencia que se tendría con una videoconferencia tradicional.

Otro sector que tampoco se ha visto afectado por la pandemia es el de los drones, tanto para uso privado, profesional como militar. Cada vez más, el uso de drones se centra en sustituir a una persona física en situaciones peligrosas para los humanos. Esto requiere que el dron posea una cámara con buena resolución y, lo que es más importante, sea capaz de transmitir la imagen de video para que se pueda operar a distancia. Otro uso muy popular es la captación de imágenes o videos aéreos. Es por esto, que el uso de cámaras 360° en drones también ha empezado a coger popularidad, aunque no se ha extendido a todos los ámbitos.

Dado que el gran punto fuerte del contenido 360° es la inmersión que ofrece, aprovecharlo en el contexto de carreras de drones ofrecería una mejor experiencia. Las carreras de drones comenzaron a realizarse como competiciones de aficionados alrededor del mundo. No fue hasta el 2015 cuando aparecieron ligas o federaciones que crearon competiciones ya más similares a al formato que se puede observar en diferentes deportes a motor como la Fórmula 1. Las carreras de drones han crecido en popularidad desde que aparecieron las primeras competiciones y, en algunos países, estas carreras se han llegado a retransmitir por televisión.

En estas carreras, donde generalmente son cuadricópteros (4 rotores) los más utilizados, los drones llevan una cámara de primera persona (FPV) en la parte frontal con la que el piloto puede observar a donde se dirige el dron. La imagen de la cámara es transmitida por el dron a unas gafas FPV que lleva el piloto que le permiten ver la imagen. En la Fig. 2 se puede observar una imagen de un dron que se utilizaría para carreras.



Fig. 2: Imagen de un dron de carreras.

Los sistemas de comunicaciones que se utilizan hoy en día en las carreras de drones no han avanzado mucho respecto a los que se utilizaban hace 6 años, a día de hoy, se sigue haciendo uso de soluciones analógicas para transmitir la señal de video. Para los estándares de comunicaciones de la actualidad, estos sistemas analógicos ofrecen muy poca capacidad de canal y una ineficiencia mayor en comparación con los sistemas digitales. Debido a esto, y junto al auge en popularidad de las carreras de drones, diversos fabricantes han actualizado las soluciones que ofrecen para incluir transmisores y receptores digitales, permitiendo la introducción de técnicas de procesado digital. Es en este contexto donde la inclusión de una cámara 360° en los drones de carreras sería una evolución coherente y mejoraría la inmersión del piloto y usuarios que visualicen el contenido.

En este trabajo se propone utilizar una cámara 360° en drones de carreras junto con el estándar MPEG-DASH para poder distribuir el contenido via streaming o VoD de la forma más eficiente. Este trabajo se realiza teniendo en cuenta que a través del canal de comunicación entre el transmisor (dron) y receptor (gafas FPV) se transmite una señal digital. Es por esto que se utilizará el canal definido en el trabajo “Algoritmos de detección en sistemas MIMO masivo para comunicaciones digitales con drones” [1]. Aquí, se propone la inclusión de una estación base (BS) MIMO (Multiple-Input Multiple-Output) masivo (mMIMO) para recibir las diferentes señales de los drones de carreras. A través de la BS se retransmite la señal a las gafas de los pilotos, ya sea de forma inalámbrica o conectadas directamente a la BS. En este caso, el trabajo se centra en la parte de la distribución de contenido desde la BS hasta los usuarios. En la Fig. 3 se puede observar un esquema de como quedaría conformado todo el sistema.

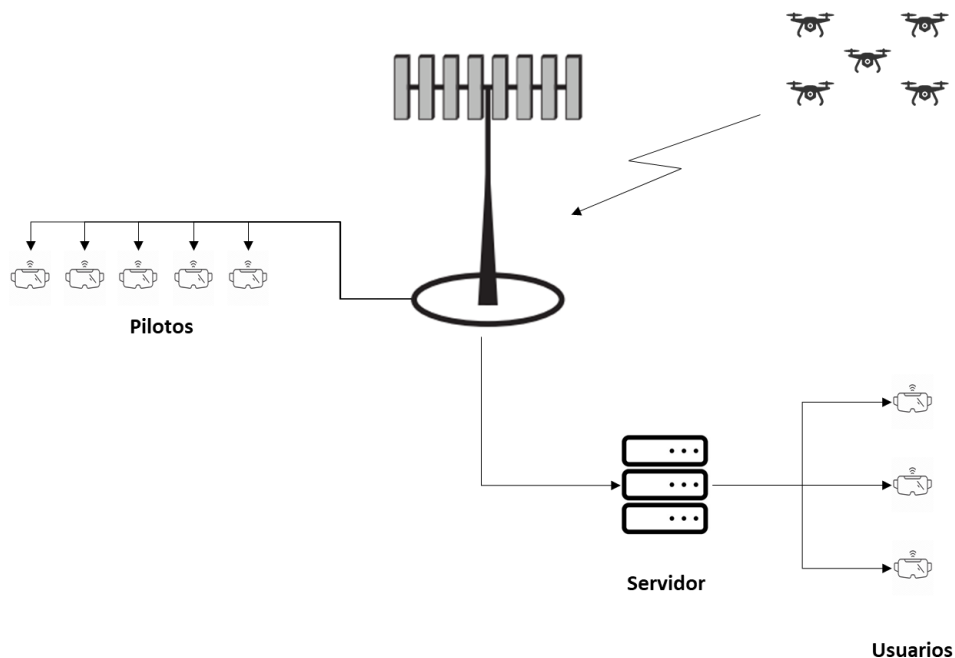


Fig. 3: Esquema de distribución de video de drones.

Con este trabajo se busca determinar si sería posible adaptar las carreras de drones para incluir sistemas de transmisiones digitales que puedan transmitir video 360° a altas calidades y transmitirlo de forma eficiente via streaming o VoD a los clientes.

II. ESTADO DEL ARTE

Tanto el consumo de videos 360° como el uso de dispositivos de RV ha aumentado y evolucionado desde que se introdujeron al gran público. Hoy en día, desde la aparición de plataformas como YouTube o Netflix, las principales formas de consumo de video son vía streaming o como VoD y priorizando la reproducción ininterrumpida a costa de la calidad del contenido.

Las carreras de drones de las principales competiciones suelen acontecer en espacios habilitados para la ocasión, generalmente cerrados. Los sistemas de comunicación que se utilizan no tienen características especiales, son dispositivos comerciales. Se utilizan dos sistemas separados para la comunicación, uno para la comunicación entre el dron y la controladora de vuelo y otro para la transmisión de video a las gafas FPV.

II.1. TRANSMISIÓN DE VIDEO EN CARRERAS DE DRONES

La transmisión de la señal de video en las carreras de drones es unidireccional, el VTX (Transmisor de vídeo) del dron transmite la señal a través de la antena y son las gafas FPV del piloto las que reciben esta señal.

La transmisión se realiza en una frecuencia de 5.8 GHz y con una potencia que varía entre los 25 mW y 600mW. En las carreras, dado que los drones vuelan muy cerca entre si, la potencia utilizada es de 25 mW para evitar interferencias. Existen un total de 10 bandas diferentes con 8 canales cada una que se pueden utilizar para la comunicación. La banda más utilizada en las carreras es la banda R dado que dispone de una mayor separación entre los canales, 37 MHz.

Existen diferentes tipos de antenas que se pueden utilizar tanto para el VTX como para el receptor en las gafas FPV. Se pueden categorizar en función de la polarización y de la directividad de la antena y, dependiendo del uso que se haga del dron, unas antenas serán una mejor opción que otras.

En una carrera de drones se conoce de antena mano en que espacio volará el dron, sin embargo, debido a su alta movilidad una antena con una gran directividad no sería eficiente ya que el dron no va a mantener la misma orientación en toda la carrera. Es por esto que en el dron las antenas omnidireccionales son las más utilizadas. En el receptor, dado que el piloto no se desliza durante la carrera, si que es posible utilizar una antena que tenga una mayor directividad. Además, debido a esta movilidad, la antena que se utiliza en el dron tiene una polarización circular. Las antenas con polarización lineal son utilizadas en el receptor en ocasiones donde el movimiento del dron no es tan rápido y es más fácil mantener la antena alineada. En la Fig. 4 se pueden observar dos de las antenas más utilizadas, siendo la antena Pagoda omnidireccional y con polarización circular y la antena Parche direccional y circular o lineal.

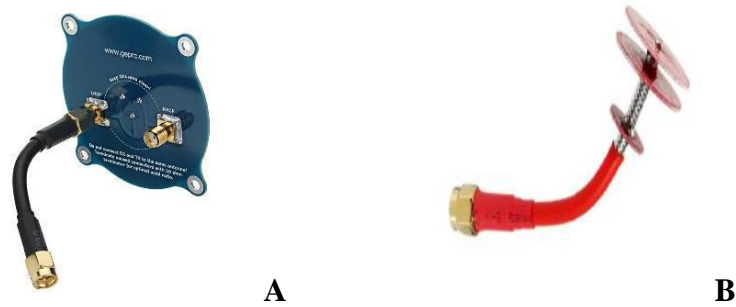


Fig. 4: Antenas Parche (A) y Pagoda (B).

II.2. HEAD-MOUNTED DISPLAY

Los HMD, popularmente referidos como cascos de RV, aparecieron como producto comercial al comienzo de la década de 2010. En la práctica, su función es idéntica a la que se podría esperar de unas gafas FPV. Sin embargo, estos han sido pensados para representar un entorno en 3 dimensiones virtual a nuestro alrededor. Con el paso de los años desde su aparición la tecnología ha ido avanzando y cada vez se han presentado nuevos dispositivos con mejores prestaciones, siendo una de ellas el campo de visión, o FoV (Field of View), característica que comparte con las gafas FPV.

Tradicionalmente, los HMD siempre se han utilizado en conjunto con un ordenador, en el que los HMD son los encargados de reproducir las imágenes generadas por el ordenador y captar la información relacionada con el movimiento del usuario. Esto siempre ha sido un factor limitante ya que, sumado al elevado coste de los HMDs, un ordenador capaz de generar estas imágenes a una tasa de fotogramas adecuada también tiene un coste elevado. En la Fig. 5 se pueden observar los componentes del HMD Valve Index, uno que requiere el uso de un ordenador para funcionar.



Fig. 5: HMD Valve Index con sus componentes.

Es por esto que la aparición del HMD de Facebook, que no requiere del uso de un ordenador para funcionar, ha sido un gran éxito. Oculus Quest 2 es un HMD que es capaz de funcionar de forma autónoma y ofrecer una experiencia similar a la de un HMD tradicional. En la Fig. 6 se puede observar una imagen de este HMD.



Fig. 6: HMD de Facebook, Oculus Quest 2.

Este HMD que es capaz de operar de forma autónoma sería un candidato ideal para su uso en el entorno de una carrera de drones, ya que no se requiere la ayuda de un ordenador externo para operar. Además, permite la conexión por cable mediante USB tipo C, lo que haría posible conectarlo a un ordenador en el caso de que fuera necesario.

II.3. *CONSUMO DE CONTENIDOS EN STREAMING Y VOD*

El consumo de contenidos digitales se ha disparado este último año a causa de la pandemia sufrida en 2020 y, en la actualidad, el consumo de estos contenidos se mantiene en niveles superiores a los que se pudieron observar en años previos a la pandemia.

Según el estudio de uso y actitudes de consumo de contenidos digitales del 2021 publicado por el Observatorio Nacional de Tecnología y Sociedad (ONTSI) [2], el consumo de servicios y contenidos digitales ha sufrido un claro incremento respecto a años anteriores siendo las películas, series, fotos, videos y música los más consumidos por la población. Como se puede observar en la Fig. 7 el 82.2% consume fotos e imágenes digitales, el 80.3% de la población consume películas, series y videos y el 80.1% consume música.

Este incremento en el consumo de servicios digitales es posible debido al aumento de las velocidades de datos. En este 2021 el 72.2% de los hogares tienen conexión por fibra óptica, un incremento considerable ya que en 2016 solo el 47.8% de los hogares lo tenían. Este incremento es importante para el consumo de video, ya que permite el consumo de contenidos con mayor resolución. Esto es observable en plataformas como YouTube o Netflix donde las máximas resoluciones admitidas han ido evolucionando con el tiempo.

El contenido inmersivo 360° lleva años abriéndose camino y, poco a poco, cada vez va aumentando su consumo. La principal desventaja de consumir contenido 360° vía streaming en contra posición al contenido tradicional (2D) es la ineficiencia a la hora de transmitir la información. En 2D la imagen que se transmite será consumida por los usuarios en su totalidad, mientras que en 360° solo se podrá visualizar en un instante el FoV de las gafas de RV que se estén utilizando. Por lo tanto, gran parte de la información que se transmite nunca será visualizada. Hoy en día, gracias al aumento de velocidad de las conexiones en los hogares es posible poder ofrecer contenido 360° vía streaming sin que la calidad de experiencia se vea afectada.

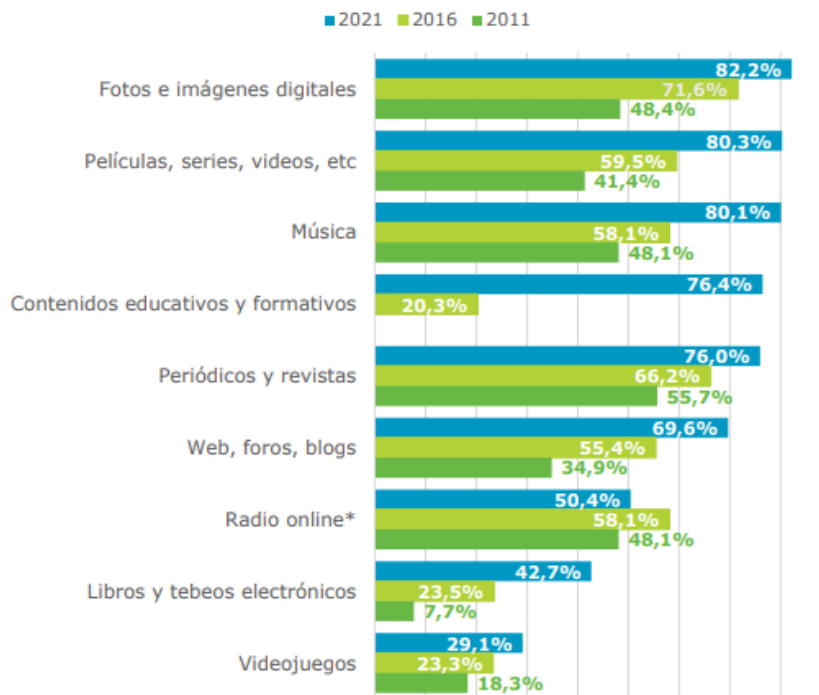


Fig. 7: Uso y consumo de contenidos digitales.

Este incremento en el consumo de servicios digitales es posible debido al aumento de las velocidades de datos. En este 2021 el 72.2% de los hogares tienen conexión por fibra óptica, un incremento considerable ya que en 2016 solo el 47.8% de los hogares lo tenían. Este incremento es importante para el consumo de video, ya que permite el consumo de contenidos con mayor resolución. Esto es observable en plataformas como YouTube o Netflix donde las máximas resoluciones admitidas han ido evolucionando con el tiempo.

La principal desventaja del contenido 360° vía streaming, en contra posición al contenido en dos dimensiones (2D), es la ineficiencia a la hora de transmitir la información. En 2D la imagen que se transmite será consumida por los usuarios en su totalidad, mientras que en 360° solo se podrá visualizar en un instante el FoV de las gafas de RV que se estén utilizando. Por lo tanto, gran parte de la información que se transmite nunca será visualizada. Hoy en día, gracias al aumento de velocidad de las conexiones en los hogares es posible poder ofrecer contenido 360° vía streaming sin que la calidad de experiencia se vea afectada.

Las cámaras 360° graban el contenido en una proyección conocida como ERP (Equirectangular Projection). En esta proyección la imagen obtenida tiene un ratio 2:1 o 16:9 y correspondería a desdoblar una esfera y dejarla extendida. Esta proyección presenta la desventaja de no tener un reparto de calidad uniforme en toda la imagen. Las zonas superiores e inferiores de la imagen (correspondientes a los polos de la esfera) tienen una mayor densidad de información que la zona central. Esto es un problema debido a que, en la gran mayoría de las situaciones, las áreas que mayor interés acaparan suelen estar entorno al centro de la imagen. En la Fig. 8 [3] se puede observar una imagen en proyección ERP, que correspondería a la totalidad de la esfera, y desde la que se extrae la imagen visualizada para el usuario a través del HMD, la cual variará en tamaño según el FoV del HMD.

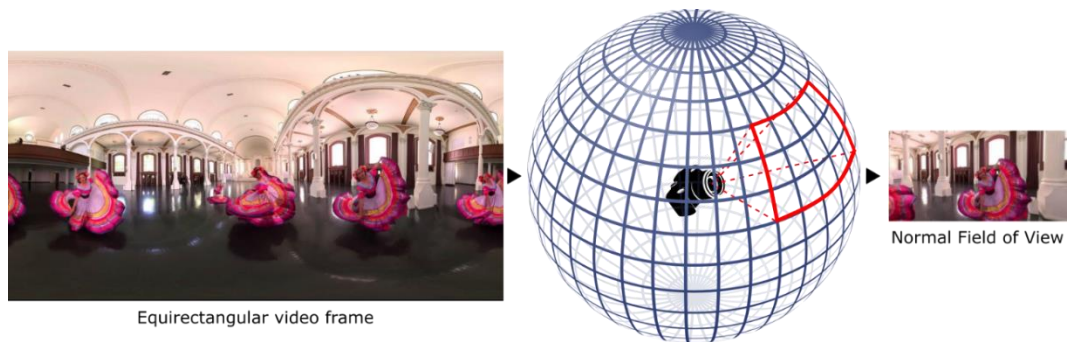


Fig. 8: Ejemplo de imagen en proyección ERP.

Debido a los problemas de la proyección ERP han surgido otras proyecciones con el objetivo de solucionar los problemas presentes en esta proyección. Una de las que más popularidad ha adquirido es la proyección CMP (Cubemap Projection). Esta proyección, en lugar de utilizar una esfera para representar la imagen hace uso de un cubo para representar la imagen, esto permite eliminar la saturación de calidad en los polos y centrar la calidad entorno al ecuador, aunque no lo hace de manera homogénea. El éxito de esta proyección se puede ver en plataformas como YouTube, donde desde hace ya tiempo se utiliza una proyección basada en CMP, EAC (Equi-Angular Cubemap) para sus videos 360°. En la Fig. 9 se puede observar como se representaría la proyección CMP en 2D.

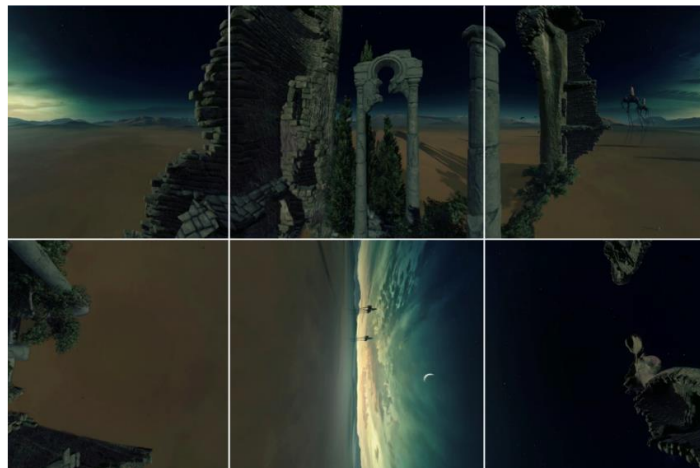


Fig. 9: Representación de la proyección CMP.

IV.4. DYNAMIC ADAPTATIVE STREAMING OVER HTTP

El streaming de video con bitrate adaptativo es una técnica que consiste en transmitir un video y, según el estado de la red, ajustar el bitrate para ofrecer la máxima calidad posible y reproducir el contenido de forma ininterrumpida. Para poder adaptar el bitrate del contenido el cliente se comunica con el servidor para, en caso de ser necesario, seleccionar una calidad diferente. Las diferentes técnicas de bitrate adaptativo funcionan sobre el protocolo HTTP. En la Fig. 10 se puede observar un esquema de como funcionaría el proceso de comunicación entre el cliente y el servidor.

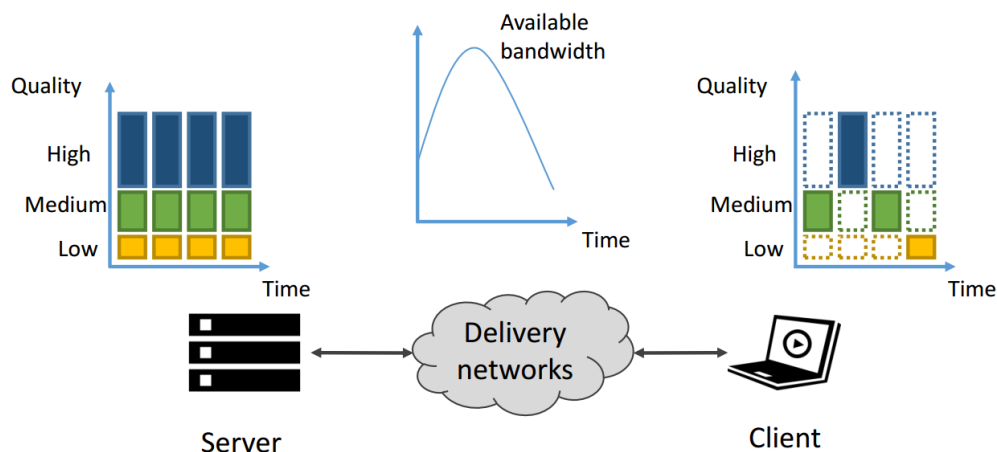


Fig. 10: Esquema de funcionamiento de técnicas de bitrate adaptativo.

MPEG-DASH es un protocolo para el streaming adaptativo de contenido multimedia estandarizado por MPEG [4]. DASH, permite el despliegue de servicios de streaming utilizando la estructura HTTP existente, lo que permite reducir los costes de despliegue. El planteamiento detrás de DASH es el de dividir un video en segmentos de corta duración y enviárselos al cliente cuando los necesite. De cada segmento se tendrán diferentes copias, cada una con un bitrate diferente, lo que permitirá cambiar de calidad simplemente seleccionando diferentes segmentos. De esta manera se consigue adaptar la calidad de la imagen al ancho de banda disponible permitiendo una reproducción ininterrumpida.

Antes de reproducir el contenido, el cliente obtiene un archivo MPD (Media Presentation Description) que es el que guarda las relaciones entre los diferentes segmentos generados, actuando como un índice de contenidos. Mediante este archivo el cliente conoce los diferentes segmentos de video almacenados en el servidor y es capaz de solicitar uno de mayor o menor calidad cuando lo necesite. En la Fig. 11 se puede observar como es estructurado el contenido dentro del archivo MPD.

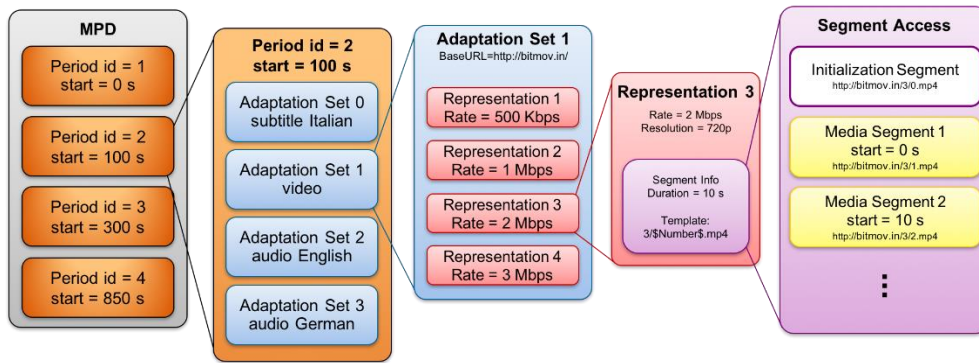


Fig. 11: Estructura de contenido en archivo MPD.

Como se puede observar en la imagen, un contenido se puede dividir en diferentes espacios de tiempo (p. ej, los episodios de una serie) lo que permite organizar el contenido e identificar en que parte de la reproducción se encuentra el reproductor. Dentro de cada Period se encuentran diferentes Adaptation Set que indican diferentes streams del contenido como audio, video o subtítulos. Para cada uno de ellos se indican diferentes representaciones (Representation), cada una con una calidad diferente. Estas representaciones son las que variarán en función del estado de la red del cliente, ofreciendo una calidad mayor o menor. Dentro de cada representación se encuentran definidos los diferentes segmentos que solicita el cliente.

III. EVALUACIÓN DE LOS COSTES DEL PROYECTO

En este apartado se van a desglosar los costes económicos y temporales derivados de la realización del proyecto.

III.1. *COSTES ECONÓMICOS*

En este trabajo se ha hecho uso de diferentes licencias de varios programas para la realización de pruebas. Se ha hecho uso del programa Matlab [5] para realizar simulaciones de canal, utilizando los módulos de comunicaciones móviles. También se ha utilizado la herramienta FFmpeg, tanto para aplicar diferentes filtros como para evaluar la calidad de un video. Estas licencias se han accedido sin ningún coste ya por ser gratuitas para estudiantes universitarios o por ser de código abierto.

Para la realización de este trabajo se ha hecho uso de un ordenador de sobremesa con las siguientes características:

- CPU: Intel core i7-6700k 4 GHz.
- GPU: Nvidia GeForce GTX 1070 8 GB.
- RAM: DDR4 2400 PC4-19200 4x8 GB.

Este ordenador ya se poseía por el alumno, por lo que no ha supuesto ningún coste su utilización. De esta forma, la realización del proyecto no ha supuesto ningún coste adicional.

III.2. *COSTES TEMPORALES*

Para la realización del proyecto se ha realizado una planificación de las tareas a realizar agrupadas por fases. En primer lugar, se ha realizado una fase de planificación donde se han definido las tareas a realizar en el proyecto además del alcance del mismo. Seguidamente se ha realizado una búsqueda del estado del arte de los aspectos referentes al proyecto para conocer las tecnologías y técnicas utilizadas en la actualidad en cuanto a transmisión de video y captura de imagen 360°. Realizadas estas tareas, se ha hecho uso del conocimiento adquirido para crear un entorno de pruebas adecuado donde realizar las diferentes simulaciones y pruebas del trabajo, como simulaciones de canal, comparación de calidades de imagen o pruebas de transmisión de video. Por último, se ha realizado la redacción de la memoria del proyecto. En la Fig. 12 se puede observar la distribución de las tareas a lo largo del tiempo, así como el coste temporal asociado a cada una de ellas.

	Fecha de inicio	Fecha fin	Duración (días)
Proyecto TFM	12/07/2021	26/11/2021	100
Planificación del proyecto	12/07/2021	20/07/2021	7
Definición de la estructura del proyecto	12/07/2021	19/07/2021	6
Definición de las tareas a realizar	12/07/2021	15/07/2021	4
Reunión con el tutor	20/07/2021	20/07/2021	1
Investigación del estado del arte	21/07/2021	11/08/2021	16
Comunicaciones en carreras drones	21/07/2021	10/08/2021	15
Transmisión video 360º	21/07/2021	03/08/2021	10
Camaras 360º	21/07/2021	22/07/2021	2
Protocolos de transmisión de video con bitrate adaptativo	21/07/2021	27/07/2021	5
Reunión con tutor	11/08/2021	11/08/2021	1
Creación de un banco de pruebas	12/08/2021	02/09/2021	16
Implementación algoritmo creación CMP 3:2	12/08/2021	25/08/2021	10
Implementación algoritmo creación CMP 11:6	12/08/2021	25/08/2021	10
Implementación algoritmo creación CMP 1:1	12/08/2021	25/08/2021	10
Implementación configuraciones CMP en player	26/08/2021	01/09/2021	5
Reunión con tutor	02/09/2021	02/09/2021	1
Realización de pruebas	03/09/2021	28/09/2021	18
Simulación de transmisión sobre canal MIMO	03/09/2021	23/09/2021	15
Simulación de reproducción de contenido	03/09/2021	09/09/2021	10
Realización de streaming de contenido	10/09/2021	13/09/2021	5
Comparación de calidad de contenido	14/09/2021	16/09/2021	5
Evaluación de pruebas realizadas	24/09/2021	27/09/2021	2
Reunión con tutor	28/09/2021	28/09/2021	1
Memoria TFM	20/09/2021	26/11/2021	50
Redacción de la memoria	20/09/2021	26/11/2021	50
Reunión con tutor	20/09/2021	23/11/2021	47

Fig. 12: Coste temporal de cada una de las tareas realizadas durante el proyecto.

IV. PROYECCIÓN CUBE-MAP PROYECTION

Como se ha comentado anteriormente, la ventaja de la configuración CMP sobre ERP es una distribución más uniforme de la calidad, además de reducir la distorsión producida en la imagen (visible cuando se observa el video en 2D) que aumenta la complejidad a la hora de comprimir el video.

Sin embargo, estas no son las únicas ventajas de utilizar una representación cúbica en lugar de una esférica. Al utilizar un cubo se tienen 6 caras claramente diferenciadas, lo que habilita que se pueda trabajar con cada una de ellas de una forma sencilla y, dado que se reduce la distorsión que se podría observar en ERP, en el caso de sustituir alguna de ellas no sería necesario tener en cuenta esta distorsión. En este trabajo se propone utilizar dos configuraciones diferentes derivadas de la proyección CMP con el objetivo de reducir la cantidad de información a enviar y manteniendo una calidad de experiencia similar respecto al contenido original.

En el contexto de una carrera de drones, aunque se haga uso de una cámara 360° para que el piloto pueda observar el entorno a su alrededor, siempre habrán zonas que tendrán un mayor tiempo de atención que otras. Un ejemplo muy claro sería comparar la cara frontal con la cara trasera o inferior, la gran mayoría del tiempo el piloto observará la cara frontal ya que así puede saber a donde se dirige el dron, mientras que las caras trasera e inferior no le aportan información relevante en la carrera. Es debido a esto que se busca reducir la calidad de las caras del cubo, a excepción de la cara frontal, para así reducir la cantidad de información a transmitir.

Las dos configuraciones propuestas parten de la representación de la proyección CMP en forma de matriz 3x2 (CMP 3:2). En la Fig. 13 se muestra como sería la disposición de las caras en esta configuración. Esta disposición no tiene porque ser absoluta, se han ordenado las caras de esta forma ya que el reproductor utilizado para esta configuración lo entiende de esta forma.

Para obtener la proyección CMP 3:2 a partir de un video en ERP, o viceversa, se ha hecho uso de la herramienta FFmpeg [6], en particular del filtro v360. FFmpeg también se ha utilizado para realizar las diferentes configuraciones.

L	F	R
T	B	Bt

Fig. 13: Representación de CMP 3:2. Caras izquierda (L), derecha (R), frontal (F), superior (T), trasera (B) e inferior (Bt).

IV.1. CONFIGURACIÓN CMP 1:1

Dado que durante una carrera de drones el piloto estará mirando al frente la gran mayoría del tiempo, esta cara es considerada como la más importante, es por esto, que toda la información de esta cara será transmitida. Además, para esta configuración también se ha tomado como caras importantes las caras laterales izquierda y derecha ya que es posible que los usuarios que visualicen el contenido, que no presentan el mismo nivel de atención que pueda presentar el piloto, quieran visualizar el entorno y observar alrededor. Con esta configuración se busca mantener la calidad en las caras que presentan un mayor interés, tanto para el piloto como para los usuarios, y reducir la calidad en las tres caras restantes: superior, trasera e inferior. El tamaño de estas caras se ha reducido a $\frac{1}{4}$ del tamaño original.

Dado que el objetivo es transmitir la menor cantidad de información posible, un aspecto a tener en cuenta es la organización de las caras para tratar de mantener una relación de aspecto lo más cercana a 1:1 posible, reduciendo el tamaño de la imagen a transmitir. Es por esto que transmitir la configuración CMP 3:2 es más eficiente que si se realizara transmitiendo las caras del cubo desplegadas en forma de T tumbada, además de que sería mucho menos eficiente ya que la mitad de la imagen que se transmite no es información útil. En la Fig. 14 se puede ver como sería este caso.

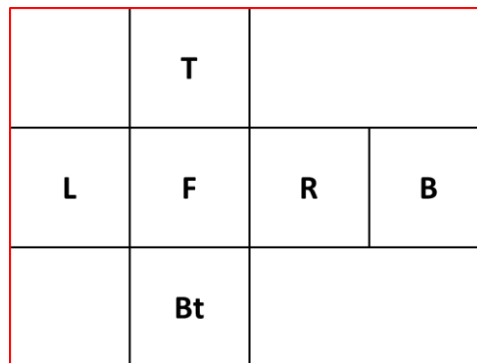


Figura 14: Imagen a transmitir con una disposición de T tumbada.

Para esta configuración se han organizado las caras lo más cercano posible a un cuadrado para mantener una relación de aspecto cercana a 1:1, ocupando el espacio restante con las caras superior, trasera e inferior. En la Fig. 15 se puede observar como quedaría la distribución de las caras en esta configuración. Esta configuración, comparada con CMP 3:2, ocupa $\frac{1}{3}$ menos y las caras central, derecha e izquierda ocupan una mayor parte de la imagen transmitida.

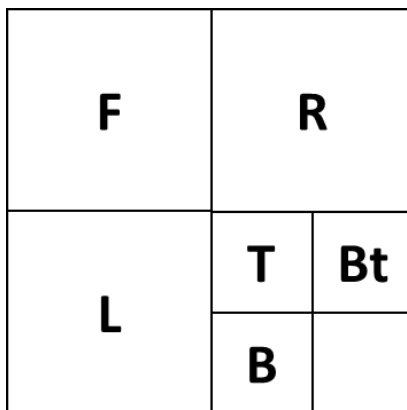


Fig. 15: Representación de la configuración CMP 1:1 propuesta.

Para generar esta configuración se han utilizado los filtros de FFmpeg para recortar, girar, redimensionar y unir las diferentes caras. Las caras frontal, derecha e izquierda no han tenido que transformarse y solo se han requerido dos cortes, uno para la cara izquierda y otro para la frontal y derecha. En el caso de las caras de menor interés, estas se han tenido que redimensionar y girar.

IV.1. CONFIGURACIÓN CMP 11:6

En la configuración propuesta anteriormente se da la misma importancia a las caras frontal, izquierda y derecha. Sin embargo, sería más acertado que la cara frontal fuera la que mayor importancia tiene y las caras laterales fueran un paso intermedio entre la cara frontal y el resto. La configuración CMP 11:6 fue propuesta en el trabajo “Conversión de formatos de proyección y codificación basada en campo de visión para streaming adaptativo de vídeos 360° utilizando DASH” [7] en el cual se proponen diferentes configuraciones de la proyección CMP. Para el contexto de una carrera de drones esta configuración puede ser muy acertada ya que crea 3 zonas diferentes dentro de la imagen: una con mucha calidad, siendo la cara frontal, otra con calidad intermedia, siendo las caras laterales, y una última con baja calidad, siendo las caras restantes. En la Fig. 16 se puede observar cual sería la disposición de las caras en esta configuración.

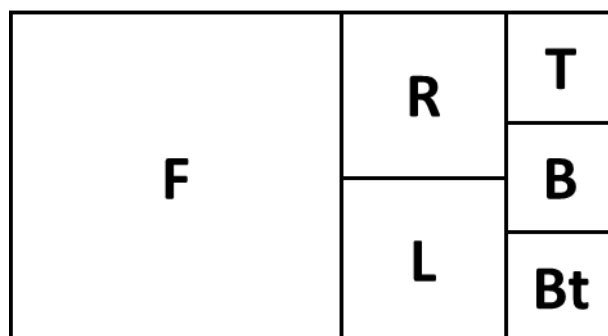


Fig. 16: Disposición de las caras del cubo en la configuración CMP 11:6.

En comparación con la configuración CMP 1:1 en esta configuración se estaría transmitiendo menos información, reduciendo así el ancho de banda necesario, pero la zona con mejor calidad queda restringida a una única cara, la frontal.

V. ESQUEMA DE DISTRIBUCIÓN DE CONTENIDO

Para distribuir el contenido a los diferentes usuarios que lo visualicen en streaming o como VoD se propone el uso del estándar MPEG-DASH para la transmisión de video con bitrate adaptativo. Dentro del ámbito de la distribución del contenido quedan enmarcados los clientes que consuman el contenido a través de Internet. Este criterio excluye a los pilotos de los drones ya que estos reciben la señal de video mediante la BS mMIMO de forma directa, sin realizar ninguna conversión a la imagen, a la cual pueden estar conectados de forma inalámbrica o mediante una conexión por cable. Haciendo uso de la posibilidad de indexar el contenido mediante el archivo MPD, es posible generar un MPD que incluya los diferentes puntos de vista con variaciones de la calidad. De esta forma un cliente podría cambiar entre diferentes streams de video dentro de la misma conexión con el servidor.

Según que salida de video se desee, ya sea en proyección CMP o ERP, se realizará un procesado diferente del video de entrada, para todos ellos se hará uso de la herramienta FFmpeg. En el caso de la proyección ERP el procesado es más simple ya que no se requiere ninguna conversión de la propia proyección. En este caso solo se realizaría la conversión del video de entrada para generar el video en formato DASH, junto con los diferentes segmentos y archivo MPD, creando varias representaciones con calidades diferentes. En la Fig. 17 se puede observar como sería un esquema de este proceso.

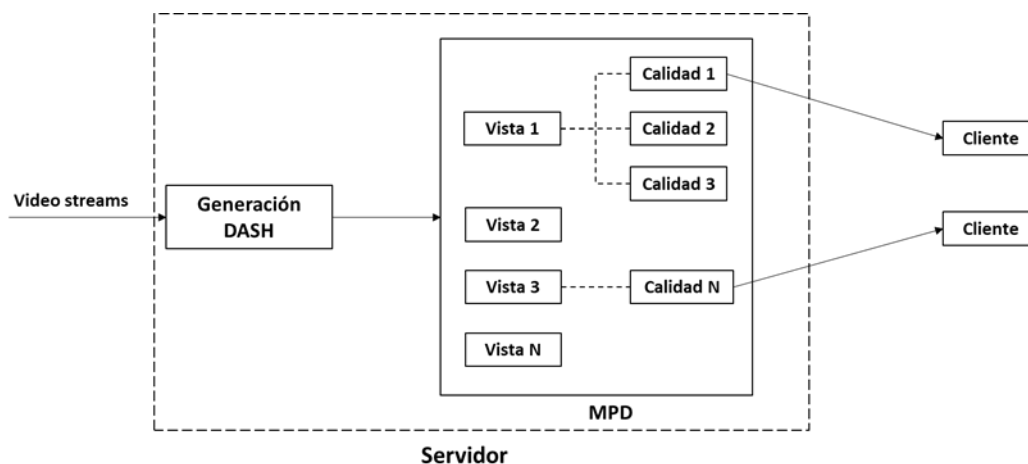


Fig. 17: Esquema de generación de stream de video en formato DASH para proyección ERP.

En el caso de la proyección CMP el proceso de generación del archivo DASH requiere pasos extra para generar esta proyección a partir de ERP. Primeramente, se debe generar la proyección CMP 3:2 a partir de ERP y, seguidamente, ajustar las caras obtenidas para generar las dos configuraciones. Una vez generada la configuración de cada stream de video entrante se procede a generar los segmentos DASH y a almacenarlos en el servidor.

Para cada configuración de CMP se han de realizar diferentes aplicaciones de filtros de FFmpeg para cada una de las diferentes caras del cubo y, una vez se han tratado cada una de las caras, se combinan cada una de ellas generando la configuración correspondiente. Aun siendo configuraciones diferentes, la cantidad de llamadas de la herramienta FFmpeg es la misma en ambos casos. En la Fig. 18 se muestra un esquema de como se realizaría este proceso en el servidor.

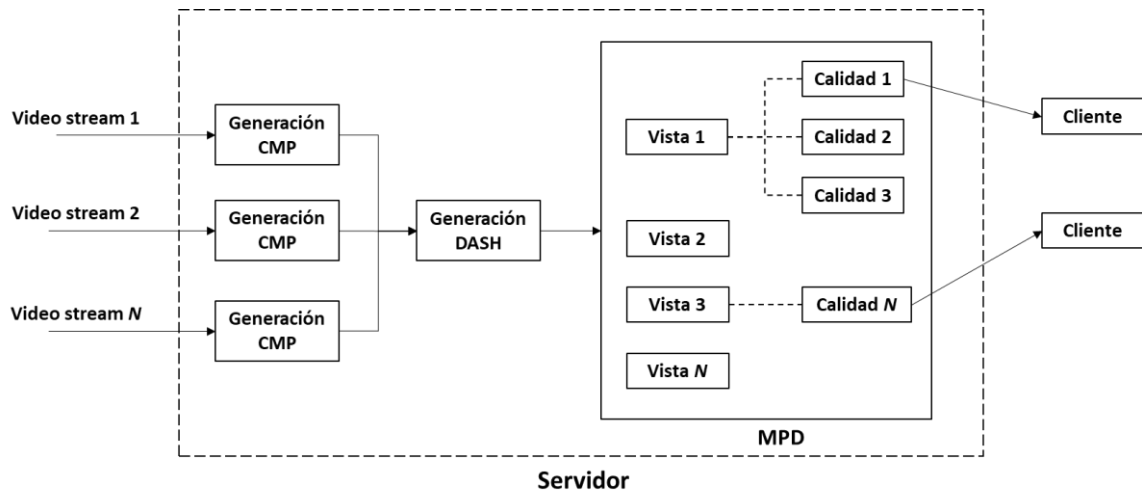


Fig. 18: Esquema de generación de stream de video DASH para configuraciones CMP.

Con este esquema propuesto se pueden tener diferentes streams de video agrupados en un solo archivo MPD, evitando que el cliente se tenga que acceder a diferentes recursos cuando quiera cambiar a un punto de vista diferente. Además, la ventaja de utilizar DASH se encuentra en la sincronización del contenido, ya que todos los streams se encuentran dentro del mismo archivo MPD. Por otro lado, sería posible combinar las diferentes configuraciones dentro de un mismo video DASH, pudiendo tener unos puntos de vista codificados con la configuración CMP 3:2, otros con CMP 11:6 y el resto con CMP 1:1. Sin embargo, al realizarse operaciones diferentes en cada configuración sería más complicado de realizar por problemas en la sincronía del contenido.

VI. PRUEBAS REALIZADAS

Para comprobar si es posible transmitir video 360° a través de un canal MIMO se ha realizado una simulación del canal utilizando el algoritmo ADMIN [8] junto con una configuración de antenas de 128x16. Se ha utilizado este caso ya que como se observó en [1], esta combinación es la que ofrecía un mejor resultado.

Para realizar esta prueba se ha realizado una simulación en Matlab, donde se ha implementado el algoritmo de detección ADMIN para un canal MIMO con modulación 64-QAM junto con OFDM. Para simular esta transmisión dentro del estándar 5G, se ha optado por utilizar un ancho de banda de 50 Mhz, con un tamaño de FFT de 5120 subportadoras, de las cuales 3010 son para señales de datos o pilotos y 2110 son nulas, las señales piloto se han incluido con una separación de 15 puntos entre cada una de ellas.

Para transmitir el video se lee fotograma a fotograma y, dado que los valores de color están comprimidos entre 0 y 255, se convierten a binario utilizando 8 bits, obteniendo un vector con todos los valores. Este vector se parte en trozos de igual tamaño que son transmitidos a través de un canal MIMO y un canal de ruido AWGN. En recepción, se aplica el detector ADMIN y, conforme se reciben las diferentes partes del fotograma, se reconstruye la imagen en el detector. En la Fig. 19 se puede un esquema de como es este proceso.

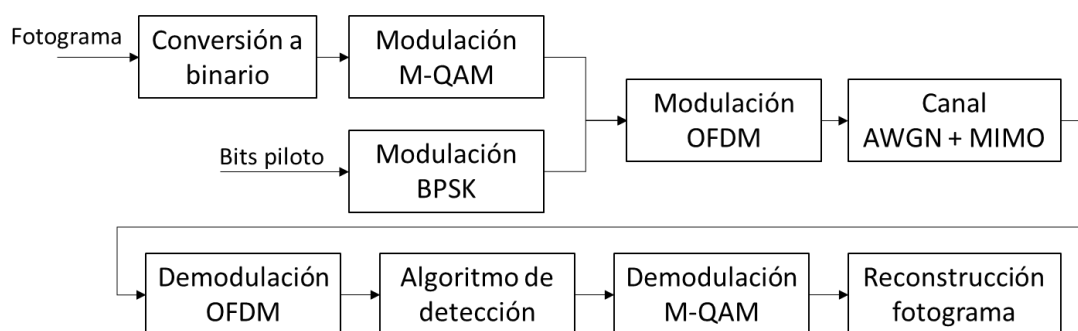


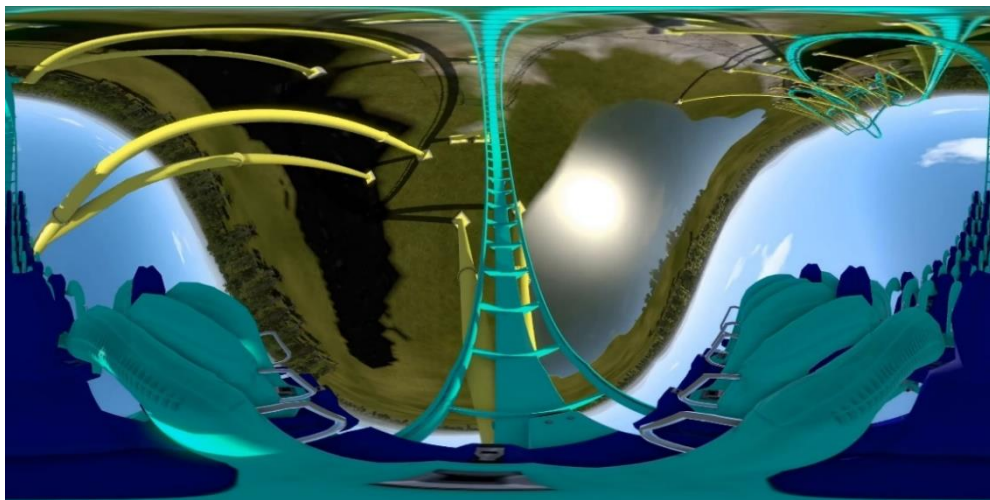
Fig. 19: Esquema de transmisión de video a través del canal MIMO.

Como muestra para realizar la simulación se ha utilizado un video con una resolución 4320x2160 (4K), 30 fotogramas por segundo y un bitrate de 50 Mbs. Se ha escogido un video con estas características para la transmisión ya que son unos requisitos que se pueden observar en soluciones comerciales para la transmisión de video para drones. En concreto estas especificaciones pertenecen a las del sistema de DJI [9].

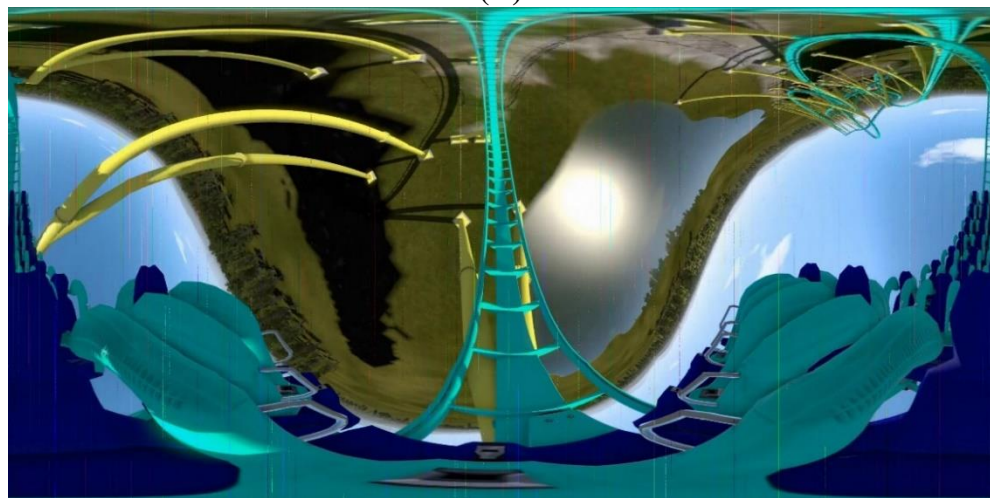
Debido a que realizar esta simulación con un video de resolución 4K es muy demandante, se ha transmitido únicamente un extracto de 2 segundos de duración para que se pudiera completar la simulación en un tiempo razonable. Dado que hoy en día no se utilizan cámaras 360° para obtener la imagen del dron que es transmitida al piloto, se ha buscado un video que muestre unas características similares a las que se observan en un video de un dron en pleno vuelo. Es por esto que se ha escogido un video 360° grabado en una montaña rusa ya que los movimientos (velocidad, giros, subidas y bajadas) son similares a los que haría un dron.

En la Fig. 20 se puede observar una comparativa entre un mismo fotograma del video original y el transmitido. Como se puede observar en la imagen, la calidad del video recibido sería una calidad aceptable para que el piloto fuera capaz de dirigir el dron sin que llegara a ser un impedimento. Para obtener una medida más objetiva de la calidad se ha hecho uso de la métrica VMAF, donde estas dos imágenes obtienen una puntuación de 90 sobre 100.

Según el resultado obtenido, la transmisión de video 360° a calidades elevadas en un canal MIMO sería posible. Sin embargo, para que este contenido pueda ser servido a clientes se tendría que incluir una parte de procesamiento digital de la señal en la transmisión, así como en el video en recepción.



(A)



(B)

Fig. 20: comparativa entre fotogramas del video original (A) y el video recibido (B).

VI.1. PRUEBA DE PLATAFORMA STREAMING

VI.1.1. PRUEBA DE REPRODUCCIÓN DE CONTENIDO VIA STREAMING

Para validar si la reproducción de contenido via streaming se realiza de forma correcta se ha realizado una transmisión del contenido adaptando los valores de resolución, fotogramas por segundo y bitrate. Para este caso solo se ha hecho uso de la proyección ERP, esto es debido a que las cámaras 360° producen el video de salida con esta misma proyección. Si se hiciera con la proyección CMP se tendría que añadir una carga extra para realizar la conversión de la imagen que, si se quiere modificar más para obtener una cierta configuración, añadiría una carga extra aun mayor. Por lo tanto, para tratar de evitar añadir mucha latencia extremo-extremo, solo se ha utilizado esta proyección.

Para realizar la prueba se ha hecho uso de una cámara 360° para generar una señal de video con la resolución, bitrate y fotogramas por segundo deseados. La señal de video es transmitida a un ordenador en el que se tiene un servidor donde se publica el contenido para que un cliente pueda acceder a el a través de un dispositivo diferente. Para el player utilizado en el cliente se ha partido como base del player de referencia de DASH-IF [10].

En el streaming con DASH se ha hecho uso de las opciones disponibles para que funcione en el modo de baja latencia (DASH-LL [11]). En este modo DASH se sirve de CMAF (Common Media Application Format) para generar “Chunks”, particiones independientes de un segmento que suelen contener unos pocos fotogramas. Con esto se consigue que el cliente no tenga que esperar a recibir todo el segmento para comenzar a reproducirlo, reduciendo los tiempos de espera. En la Fig. 21 se puede observar una comparativa sin y con CMAF. En el caso de no utilizar CMAF el player está obligado a recibir el segmento en su totalidad antes de poder empezar a reproducirlo, por el contrario, si se utiliza CMAF se puede comenzar en cuanto se reciben los primeros chunks, reduciendo el tiempo de espera.

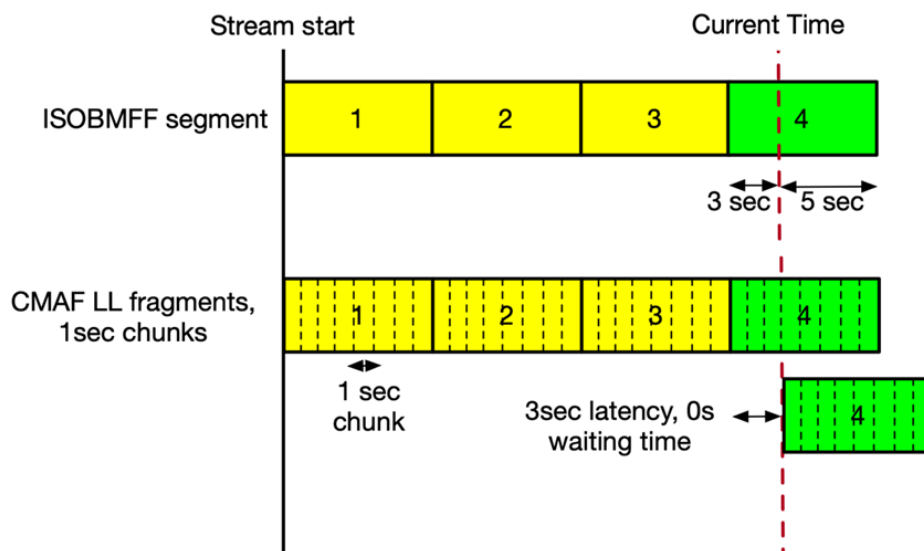
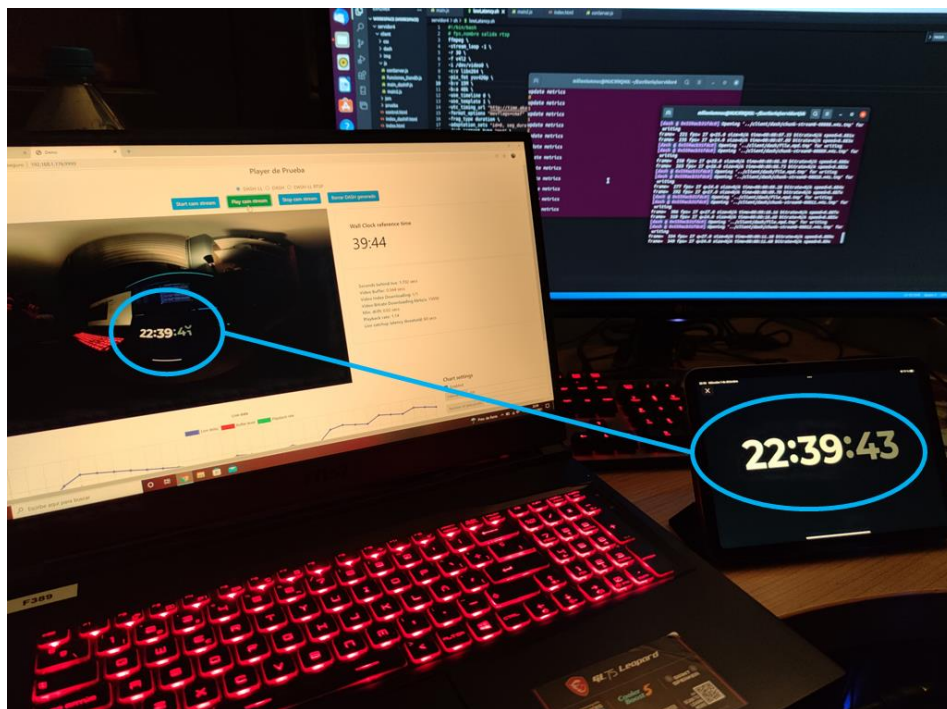


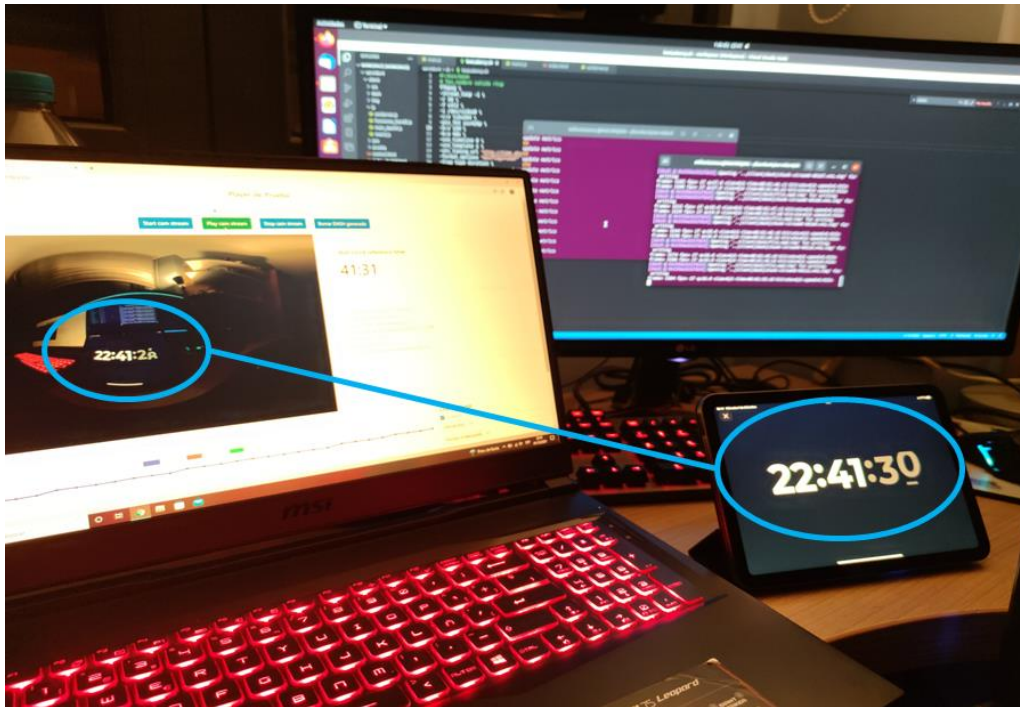
Fig. 21: Comparación de latencia entre sin y con DASH-LL.

La prueba se ha realizado con una transmisión de video en resolución 4K, 30 fotogramas por segundo y con un bitrate de 15 Mbs. En cuanto a las opciones de codificación DASH, se ha indicado una duración de segmento de 1 segundo y de 0.2 segundos para los chunks. El tamaño del GoP (Group of Picture) se ha indicado de 6 fotogramas, de esta forma, el tamaño de GoP coincide con la duración de cada chunk, haciéndolos independientes.

El bitrate se ha reducido en comparación con las soluciones comerciales debido a que la reproducción del contenido no se podía sostener de forma ininterrumpida con valores de bitrate más elevados. En la Fig. 22 se pueden observar 2 imágenes donde se puede ver una comparativa de latencia extremo-extremo realizada mientras la cámara graba un reloj. Como se puede ver en las imágenes la latencia que se obtiene es de entre 1 y 2 segundos.



(A)



(B)

Fig. 22: Capturas de la transmisión de video de latencia extremo-extremo.

El objetivo de realizar la transmisión via streaming es conseguir mantener la mejor calidad posible con la menor latencia posible para acercarse lo máximo posible a una reproducción en tiempo real. Como se ha podido ver es posible que los clientes consuman contenido de alta resolución a una baja latencia.

VI.1.2. PRUEBA DE REPRODUCCIÓN DE CONTENIDO BAJO DEMANDA

Para realizar pruebas de reproducción de video se ha hecho uso de la plataforma utilizada en [4], adaptándola para incluir una nueva configuración de la proyección CMP. En este sistema se tiene integrado la parte del servidor, donde se almacena y procesa el contenido recibido, y del cliente, donde se tiene un player compatible con MPEG-DASH.

A partir de un video en CMP 3:2 se han generado las configuraciones CMP 11:6 y CMP 1:1 que posteriormente han sido convertidas al formato DASH. Para cada video se han generado 3 calidades diferentes: Una calidad alta (50 Mbs), una calidad media (25 Mbs) y una calidad baja (5 Mbs). Dado que en las carreras de drones solo se transmite una señal de video, el audio no se ha tenido en cuenta. La prueba ha consistido en reproducir el contenido desde otro dispositivo conectado de forma inalámbrica a la misma red en la que se encuentra el servidor de contenidos.

Al reproducir los contenidos se han obtenido métricas referentes al nivel del buffer, la calidad reproducida y el throughput. En la Fig. 23 se puede observar una gráfica donde se compara la calidad seleccionada en cada instante, a lo largo de la reproducción del video. Los valores de calidad pueden ser 0, 1 o 2, siendo 2 la máxima calidad disponible. Viendo la gráfica se puede observar que en el caso de las configuraciones CMP 11:6 y CMP 1:1 la calidad en ambos casos se mantiene similar durante toda la reproducción. Sin embargo, la configuración CMP 3:2 se muestra más inestable y no mantiene tanto tiempo la máxima calidad como las otras dos configuraciones.

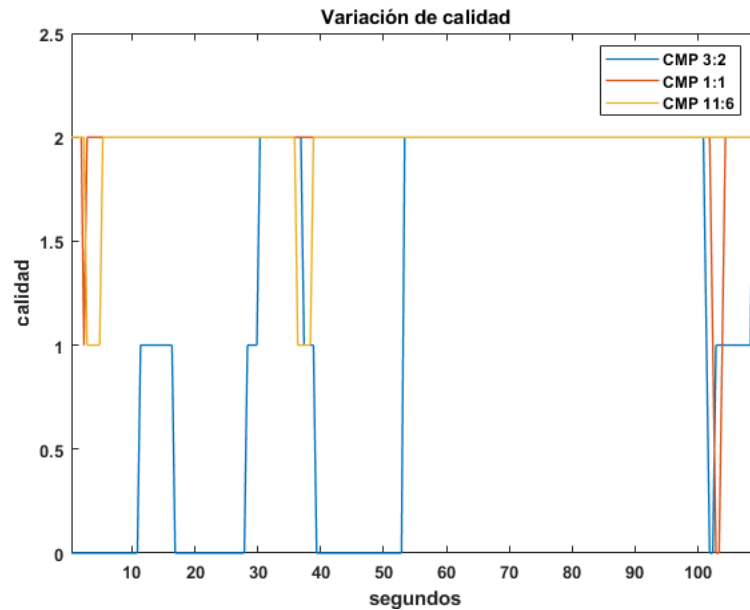


Fig. 23: Gráfica comparativa de calidad de las configuraciones CMP.

Por otro lado, en la Fig. 24 se puede observar una gráfica donde se comparan los niveles del buffer a lo largo de la reproducción del contenido de las diferentes configuraciones. Como se puede observar, las tres configuraciones presentan un nivel de en torno a 12 segundos durante la reproducción. Además, se puede apreciar claramente como, al principio de la reproducción, el buffer comienza a llenarse, se mantiene en torno a un valor estable mientras dura la reproducción del contenido y, al final de la reproducción, comienza a vaciarse consumiendo los segundos de video almacenados.

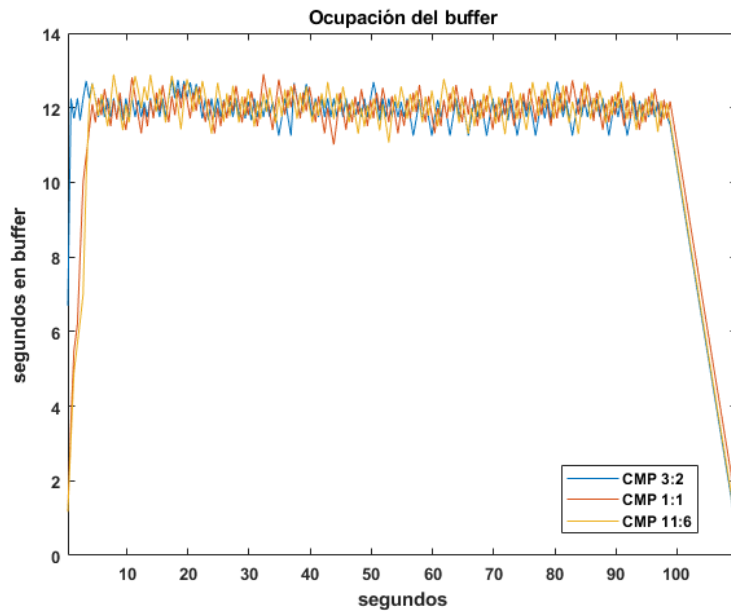


Fig. 24: Gráfica comparativa del nivel de buffer de las configuraciones CMP.

Por último, en la Fig. 25 se muestra la gráfica obtenida al comparar el throughput de las 3 configuraciones durante la reproducción del video. Como se puede observar en la gráfica las configuraciones CMP 11:6 y CMP 1:1 mantienen un throughput similar durante toda la reproducción, en torno a 0.25 MB/s. La configuración CMP 3:2 mantiene un valor de throughput un poco inferior a las otras dos configuraciones, en especial se puede observar esta diferencia al principio de la reproducción del contenido. Esta diferencia en el throughput se puede entender ya que la configuración CMP 3:2 es la que varía más la calidad durante la reproducción y la que se encuentra más tiempo en las calidades más bajas.

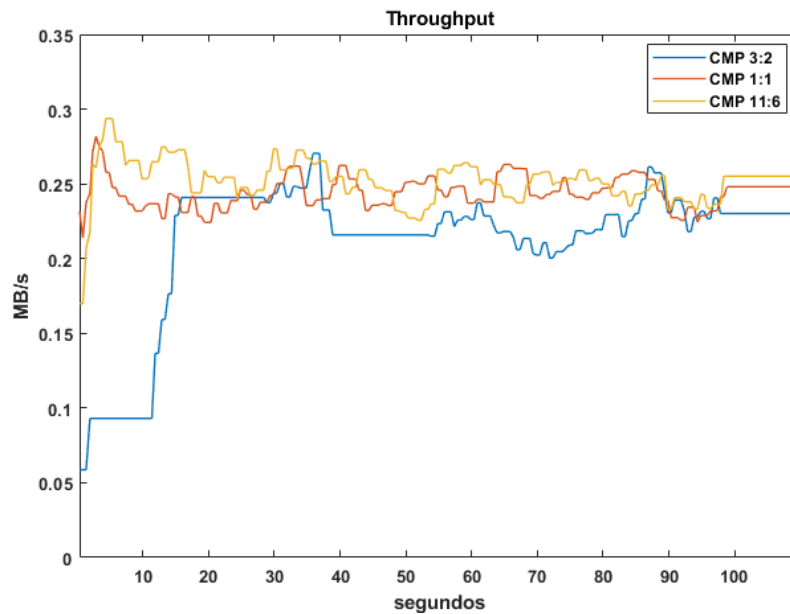


Fig. 25: Gráfica comparativa del throughput de las configuraciones CMP.

VI.2. *CONCLUSIONES DE PRUEBAS REALIZADAS*

Tras la realización de las pruebas se ha podido observar que la inclusión de una plataforma de streaming de video 360° mediante DASH-LL ofrece un resultado aceptable desde el punto de vista de la latencia extremo-extremo. En cuanto a la reproducción de contenido VoD con configuraciones de la proyección CMP, se observa que se puede obtener un mejor resultado con las configuraciones CMP 1:1 y CMP 11:6 en comparación con CMP 3:2, además de las ventajas añadidas de utilizar la proyección CMP en lugar de ERP.

VII. CONCLUSIÓN Y LÍNEAS DE TRABAJO FUTURAS

En este trabajo se ha estudiado la inclusión de cámaras 360° en el contexto de las carreras de drones y la distribución del contenido mediante streaming o como VoD. Este escenario se ha propuesto utilizando una BS mMIMO para recibir la señal de video transmitida por los drones y distribuirla a los pilotos y a un servidor de contenidos.

Con los resultados obtenidos se ha podido ver que sería posible realizar la transmisión de la señal de video de los drones durante una carrera para trasmitirla a diferentes usuarios. Con el uso de técnicas de bitrate adaptativo, DASH en este caso, se consigue una reproducción ininterrumpida manteniendo una latencia de entre 1 y 2 segundos extremo-extremo. Además, se ha podido comprobar que es posible reproducir el contenido como VoD manteniendo una calidad similar al video original utilizando diferentes configuraciones de la proyección CMP. De esta manera se consigue concentrar la calidad de la imagen en las zonas de interés, reduciendo el ancho de banda necesario.

VII.1. LÍNEAS DE TRABAJO FUTURO

Aunque en este proyecto se ha visto la viabilidad de transmisión de video 360° junto con mMIMO en el contexto de una carrera de drones, es posible seguir investigando en diferentes aspectos para mejorar la calidad del contenido ofrecido.

Por un lado, se puede realizar una investigación comparando proyecciones diferentes para la transmisión de la imagen de video. A parte de las proyecciones ERP y CMP existen otras proyecciones basadas en diferentes poliedros que permiten representar las imágenes en 3 dimensiones. Como se ha comentado, CMP mejora a ERP manteniendo las zonas de mayor calidad entorno al ecuador de la imagen, sin embargo, la calidad no es homogénea en esa zona. Utilizar una proyección con un reparto homogéneo de la calidad podría permitir una mejor QoE para los usuarios.

Por otro lado, dado que las cámaras 360° suelen grabar en proyección ERP, la conversión a otra proyección añade un coste adicional de tiempo, repercutiendo en la latencia. Es por esto que una mejora sería la inclusión de zonas de interés en la proyección ERP que se actualizarán en función de la visión del usuario. Esto permitiría reducir la cantidad de información a transmitir y centrar la calidad en la zona de interés. En la Fig. 26 se puede observar un ejemplo de la inclusión de una zona de interés en un video, donde se mantiene la calidad en la zona alrededor del personaje y el resto de la imagen se muestra en una menor calidad.



Fig. 26: Ejemplo de la aplicación de una zona de interés en un video.

Por último, las cámaras incluidas en los drones para capturar la imagen no están posicionadas con la vista paralela al suelo, se colocan en el dron con un cierto ángulo de inclinación para contrarrestar la del dron cuando acelera. Sería de interés poder conocer este ángulo para que la generación del cubo o esfera tuviera este aspecto en cuenta y se pudiera generar la proyección de forma más correcta.

REFERENCIAS

- [1] Fernández Dasí, M. (2021). Algoritmos de detección en sistemas MIMO masivo para comunicaciones digitales con drones. Universitat Politècnica de València. <http://hdl.handle.net/10251/174484>
- [2] <https://www.mineco.gob.es/portal/site/mineco/menuitem.d27e450d6789dd5c6a5af299026041a0/?vgnnextoid=6863cebeb6ceb510VgnVCM1000001d04140aRCRD>
- [3] <http://blog.nitishmutha.com/equirectangular/360degree/2017/06/12/How-to-project-Equirectangular-image-to-rectilinear-view.html>
- [4] <https://mpeg.chiariglione.org/standards/mpeg-dash>
- [5] <https://es.mathworks.com/products/matlab.html>
- [6] <https://www.ffmpeg.org/>
- [7] Conversión de formatos de proyección y codificación basada en campo de vision para streaming adaptativo de videos 360° utilizando <https://roderic.uv.es/handle/10550/80930>
- [8] S. Shahabuddin, M. Juntti and C. Studer, "ADMM-based infinity norm detection for large MU-MIMO: Algorithm and VLSI architecture," 2017 IEEE International Symposium on Circuits and Systems (ISCAS), 2017, pp. 1-4, doi: 10.1109/ISCAS.2017.8050311.
- [9] <https://www.dji.com/es/dji-fpv/specs>
- [10] <https://reference.dashif.org/dash.js/latest/samples/low-latency/index.html>
- [11] <https://github.com/Dash-Industry-Forum/dash.js/wiki/Low-Latency-streaming>