UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Microelectrónica Aplicada
Sistemas de información y Comunicaciones

Máster en
Tecnologías de Telecomunicación

Trabajo Fin de Máster

Estudio y caracterización del estándar de compresión de vídeo HEVC

Autor: Nayra Hernández Medina
Tutor(es): Dr. D. Sebastián López Suárez
Dr. D. Roberto Sarmiento Rodríguez
Fecha: Diciembre 2013

IUMA
t +34 928 451 086
f +34 928 451 083
iuma@iuma.ulpgc.es
www.iuma.ulpgc.es
Campus Universitario de Tafira
35017 Las Palmas de Gran Canaria

Versión: 2.0 – 14 junio 2011

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Microelectrónica Aplicada
Sistemas de información y Comunicaciones

Máster en
Tecnologías de Telecomunicación

Trabajo Fin de Máster

Estudio y caracterización del estándar de compresión de vídeo HEVC

HOJA DE FIRMAS

Alumno/a:	Nayra Hernández Medina	Fdo.:
Tutor/a:	Dr. D. Sebastián López Suárez	Fdo.:
Tutor/a:	Dr. D. Roberto Sarmiento Rodríguez	Fdo.:

Fecha: Diciembre 2013

IUMA
t +34 928 451 086
f +34 928 451 083
iuma@iuma.ulpgc.es
www.iuma.ulpgc.es
Campus Universitario de Tafira
35017 Las Palmas de Gran Canaria

Versión: 2.0 – 14 junio 2011

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Microelectrónica Aplicada
Sistemas de información y Comunicaciones

Máster en
Tecnologías de Telecomunicación

Trabajo Fin de Máster

Estudio y caracterización del estándar de compresión de vídeo HEVC

HOJA DE EVALUACIÓN

Calificación: .....

Presidente	Dr. D. Juan A. Montiel	Fdo.:
Secretario	Dr. D. Luis Hernández Acosta	Fdo.:
Vocal	Dr. D. José Ramón Sendra Sendra	Fdo.:

Fecha: Diciembre 2013

IUMA
t +34 928 451 086
f +34 928 451 083
iuma@iuma.ulpgc.es
www.iuma.ulpgc.es
Campus Universitario de Tafira
35017 Las Palmas de Gran Canaria

Índice

CAPÍTULO 1: ANTECEDENTES Y OBJETIVOS 12
1.1. INTRODUCCIÓN 12
1.2. ANTECEDENTES 13
1.3. OBJETIVOS 19
1.4. ORGANIZACIÓN DE LA MEMORIA 20
CAPÍTULO 2: ESTÁNDAR DE COMPRESIÓN DE VÍDEO HEVC 21
2.1. ESTANDARIZACIÓN 21
2.2. TECNOLOGÍA Y CARACTERÍSTICAS CLAVE 23
2.2.1. CODING TREE UNITS Y CODING TREE BLOCK STRUCTURE 24
2.2.2. TRANSFORM UNITS Y TRANSFORM BLOCKS 26
2.2.3. PREDICCIÓN ESPACIAL (INTRA) 27
2.2.4. PREDICCIÓN TEMPORAL (INTER) 28
2.2.5. CODIFICACIÓN ENTRÓPICA 31
2.2.6. FILTROS IN-LOOP 32
2.2.7. PERFILES 34
CAPÍTULO 3: CÓDIGO DE REFERENCIA 36
3.1. INTRODUCCIÓN 36
3.2. MÓDULOS PRINCIPALES 37
3.3. PARÁMETROS PARA LA CODIFICACIÓN 40
CAPÍTULO 4: ANÁLISIS FUNCIONAL 43
4.1. INTRODUCCIÓN 43
4.2. CONDICIONES DE TEST 44
4.3. SECUENCIAS DE TEST 46
4.4. PLATAFORMA DE TEST 47
CAPÍTULO 5: RESULTADOS Y CONCLUSIONES 48
5.1. INTRODUCCIÓN 48
5.2. RESULTADOS 49
5.3. CONCLUSIONES 57
5.4. MERCADO 58
BIBLIOGRAFÍA 60

Índice de Figuras

FIGURA 1 COMPARACIÓN DE LAS PRESTACIONES DEL ESTÁNDAR H.264/AVC CON LOS ESTÁNDARES DE COMPRESIÓN MPEG-4, MPEG-2 Y H.263 CODIFICANDO LA SECUENCIA "TEMPETE" A 30 FOTOGRAMAS POR SEGUNDO 17

FIGURA 2 COMPARACIÓN DE LAS PRESTACIONES DEL ESTÁNDAR HEVC CON ESTÁNDARES PREVIOS CODIFICANDO LA SECUENCIA "JOHNNY" A 60 FOTOGRAMAS POR SEGUNDO 18

FIGURA 3 CODIFICADOR DE VÍDEO HÍBRIDO PARA HEVC [8] 23

FIGURA 4 PARTICIONADO EN H.264 (IZQUIERDA) Y EN HEVC/H.265 (DERECHA) 25

FIGURA 5 MODOS DE DIVISIÓN DE UN CB EN PBS (SUJETO A RESTRICCIONES). PARA INTRA, SÓLO SON SOPORTADAS LAS DIVISIONES MXM Y M/2XM/2 26

FIGURA 6 DIVISIÓN DE UN CTB EN CBS (LÍNEAS SÓLIDAS) Y TBS (LÍNEAS DISCONTINUAS) (IZQUIERDA) Y SU CORRESPONDIENTE QUADTREE (DERECHA) 27

FIGURA 7 MODOS Y ORIENTACIONES DIRECCIONALES EMPLEADOS EN LA PREDICCIÓN ESPACIAL PARA HEVC (IZQUIERDA) Y H.264 (DERECHA) 28

FIGURA 8 MUESTRAS EN POSICIONES ENTERAS Y FRACCIONALES PARA LA INTERPOLACIÓN DE LA LUMA 29

FIGURA 9 MÉTODOS DE ESCANEO EN HEVC. (A) DIAGONAL UP-RIGHT. (B) HORIZONTAL. (C) VERTICAL 32

FIGURA 10 CUATRO GRADIENTES USADOS EN SAO MODO "EDGE OFFSET". (A) HORIZONTAL. (B) VERTICAL. (C)135° DIAGONAL . (D) 45° DIAGONAL 33

FIGURA 11 ESTRUCTURA DE CODIFICACIÓN JERÁRQUICA DEL MODO DE CONFIGURACIÓN RA 44

FIGURA 12 ESTRUCTURA DE CODIFICACIÓN JERÁRQUICA DEL MODO DE CONFIGURACIÓN LB 45

FIGURA 13 RESULTADOS PARA LA SECUENCIA CLAIRE 51

FIGURA 14 RESULTADOS PARA LA SECUENCIA MOBILE 51

FIGURA 15 RESULTADOS PARA LA SECUENCIA BLOWINGBUBBLES 52

FIGURA 16 RESULTADOS PARA LA SECUENCIA RACEHORSES 52

Índice de Tablas

TABLA 1 HITOS IMPORTANTES EN EL PROCESO DE ESTANDARIZACIÓN 22

TABLA 2 COEFICIENTES EMPLEADOS PARA LOS FILTROS DE 8 Y 7 ETAPAS PARA LA LUMA, MEDIO PÍXEL (HFILTER) Y CUARTO DE PÍXEL (QFILTER), RESPECTIVAMENTE 30

TABLA 3 COEFICIENTES EMPLEADOS PARA LOS FILTROS DE 4 ETAPAS PARA LAS MUESTRAS FRACCIONALES DE CROMA 30

TABLA 4 EDGEIDX CATEGORÍAS PARA SAO EN MODO "EDGE OFFSET" 33

TABLA 5 PERFILES PARA HEVC 34

TABLA 6 COMPARATIVA ENTRE HEVC MAIN STILL PICTURE Y JPEG 2000 Y JPEG EN CUANTO A TASA DE BIT BASÁNDOSE EN PSNR (PEAK SIGNAL TO NOISE RATIO) Y MOS (MEAN OF SCORE) 35

TABLA 7 MÓDULOS DEL HM 8.0 37

TABLA 8 CLASES DEL HM 8.0 38

TABLA 9 FICHEROS DEL HM 8.0 39

TABLA 10 PARÁMETROS PRINCIPALES EN LA CODIFICACIÓN 42

TABLA 11 LISTA DE SECUENCIAS DE TEST 46

TABLA 12 BIT RATE Y PSNR OBTENIDOS EN AI 49

TABLA 13 BIT RATE Y PSNR OBTENIDOS EN RA 49

TABLA 14 BIT RATE Y PSNR OBTENIDOS EN LB 50

TABLA 15 BIT RATE Y PSNR OBTENIDOS EN LP 50

TABLA 16 DISTRIBUCIÓN DE COMPLEJIDAD PARA ALL INTRA 54

TABLA 17 DISTRIBUCIÓN DE COMPLEJIDAD PARA RANDOM ACCESS 54

TABLA 18 DISTRIBUCIÓN DE COMPLEJIDAD PARA LOW DELAY B 55

TABLA 19 DISTRIBUCIÓN DE COMPLEJIDAD PARA LOW DELAY P 55

TABLA 20 MEDIA DE LAS FASES DE CODIFICACIÓN CON MAYOR COMPLEJIDAD 56

TABLA 21 IMPLEMENTACIONES Y PRODUCTOS HEVC 59

Capítulo 1: Antecedentes y Objetivos

Es indiscutible el increíble auge de las aplicaciones multimedia que está experimentando la sociedad actual. Cada vez son más los usuarios que demandan algún tipo de servicio multimedia, tales como videoconferencia, transmisión y recepción de fotos o secuencias de vídeo… Debido a que la mayoría de las aplicaciones multimedia que se van incorporando ofrecen cada vez más servicios a los usuarios a costa de resultar complejas, la transmisión y recepción de toda la información que se requiere para realizar los servicios con una calidad aceptable resulta ser una tarea de elevada dificultad. Es por esto que resulta indispensable crear una serie de estándares que permitan optimizar la codificación y decodificación de la información transmitida y de manera muy especial, la información de vídeo.

1.1. Introducción

Es indiscutible la importancia que está adquiriendo en el mercado actual la demanda de productos multimedia. Cada vez son más los usuarios que demandan algún tipo de servicio multimedia, tales como transmisión y recepción de fotos o vídeos o videoconferencias. Debido a que la información que se desea mandar o recibir es elevada, de manera especial en la información de vídeo, resulta de gran interés el desarrollo de estándares que permitan codificar y decodificar la información conservando la calidad de la misma.

La meta de este trabajo es estudiar y caracterizar el nuevo estándar de codificación de vídeo High Efficiency Video Coding (HEVC). Este nuevo estándar necesitará la mitad de la velocidad de bits de su antecesor el ITU-T H.264 y como ya hiciera éste último, se prevé que esta nueva norma impulsará un nuevo ciclo de innovación en la industria relacionada con la codificación de vídeo. Esta industria evoluciona rápidamente, es por ello, por lo que el nuevo estándar ofrecerá una solución flexible, fiable y robusta tal que servirá para los próximos 10 años.

Con el HEVC se inicia una fase de innovación en la producción de vídeos que abarca todo el ámbito de las TIC.

1.2. Antecedentes

La principal meta de todos los estándares de codificación de vídeo ha sido optimizar la eficiencia de la codificación. Con eficiencia de la codificación se hace referencia a la habilidad de minimizar la tasa de bit necesaria para obtener un nivel de calidad de vídeo específica, o formulado de otra manera, para maximizar la calidad de vídeo con una tasa de bit dada. Los estándares de codificación de vídeo han evolucionado principalmente a través de los estándares ITU-T e ISO/IEC. A continuación se presenta brevemente, en orden cronológico, cada uno de los estándares de codificación previos al HEVC.

- H.261

  El primer estándar de codificación de vídeo de la ITU-T que tuvo un impacto significativo, denominado H.261 [1], fue desarrollado entre finales de los 80 y principios de los 90 para la aplicación particular de videoconferencia (comunicación *full-dúplex* mediante un enlace de vídeo) en RDSI (Red Digital de Servicios Integrados) de banda estrecha. En este sentido, H.261 alcanza sus máximas prestaciones para anchos de banda múltiplos enteros de 64 Kbps.

- MPEG-1

  A principios de los 90, ISO/IEC estandarizó su primer estándar, el MPEG-1 [2]. Se diseñó para proporcionar compresión de vídeo y audio para almacenamiento y reproducción en CD-ROMs. El objetivo de MPEG-1 es el de comprimir vídeo y audio a una tasa de bit de 1.5 Mbps con una calidad comparable a las cintas de vídeo VHS. El vídeo CD nunca tuvo éxito comercial, ya que la mejora de calidad con respecto a VHS no era suficiente para tentar a los consumidores a reemplazar sus grabadores de vídeo. Sin embargo, MPEG-1 es importante por dos razones: ha ganado un uso extenso en otras aplicaciones de almacenamiento y transmisión de vídeo y, además, su funcionalidad se ha usado y extendido en el popular estándar H.262/MPEG-2.

- H.262/MEPG-2

  El lanzamiento de MPEG-1 tuvo mucho éxito y surgieron múltiples productos en el mercado basados en él. Un factor clave de este éxito es la estructura genérica del estándar y el amplio rango de parámetros específicos aplicables. Basado en este éxito, ITU-T e ISO/IEC se unieron para entre 1995 y 1996 presentar H.262/MPEG-2 [3]. Dicha fase pretende extender las características de la primera, manteniendo con ésta una compatibilidad hacia atrás, con el principal objetivo de obtener un estándar para la codificación de imágenes de televisión, con buena calidad, usando una tasa de transmisión por debajo de 10 Mbits/s. Posteriormente, el alcance de MPEG-2 se amplió a la codificación de televisión de alta definición (*HDTV- High Definition Television*), proyecto que inicialmente fue denominado MPEG-3, hasta resultar finalmente integrado en MPEG-2. Este esfuerzo de estandarización realizado en esta segunda fase de MPEG fue aprovechado por aplicaciones emergentes como la distribución de televisión por cable, almacenamiento digital de alta capacidad en DVD (*Digital Versatile Disc*), los servicios de bases de datos en red usando ATM y la difusión de vídeo terrestre y vía satélite. Para lograr este objetivo, MPEG-2 conserva las técnicas de compresión más eficaces dentro del estándar MPEG-1 como la predicción bidireccional y la estimación de movimiento con precisión de medio píxel. Además, en MPEG-2 se añaden nuevas características para lograr codificar no sólo imágenes progresivas (no-entrelazadas) sino también entrelazadas (las que se utilizan en las transmisiones de televisión).

- MPEG-4

  La segunda parte del estándar MPEG tuvo todavía más éxito que su predecesora, convirtiéndose en el estándar para la televisión digital, tanto en su difusión por canales digitales de banda ancha, como en su almacenamiento sobre soporte digital (DVD). Sin embargo, la filosofía básica de las dos primeras fases era la misma, codificar eficientemente una sucesión de imágenes acompañadas de audio intentando emular las características básicas de un reproductor de vídeo analógico. Tras un intento fallido de desarrollar una tercera parte del estándar MPEG destinada a la codificación de HDTV, que terminó siendo incluida en MPEG-2 como se ha mencionado, los investigadores del grupo ISO/IEC emprendieron un nuevo proyecto que dio lugar al estándar MPEG-4 [4]. El estándar MPEG-4 se

desarrolló con el propósito de extender las capacidades de los otros dos estándares para tasas de transmisión bajas, además de proporcionar interactividad con el usuario. Para lograr este objetivo, este estándar contempla el uso de técnicas avanzadas de compresión además de dividir la secuencia en objetos audiovisuales, permitiendo de esta manera el acceso y manipulación de su contenido.

- H.263

  De manera paralela al desarrollo del estándar MPEG-4, el grupo VCEG de la ITU comenzó el desarrollo del estándar H.263 [5]. La finalidad de este estándar no es otra que, a partir de la estructura básica del estándar H.261, mejorar las prestaciones de compresión para tasas de compresión muy bajas, típicamente por debajo de 64 Kbps. Para conseguir este objetivo, el estándar define un núcleo básico de compresión similar al especificado en el estándar H.261 y un conjunto de 18 modos avanzados de compresión opcionales en forma de anexos al estándar. La primera versión de H.263 introduce cuatro modos opcionales, cada uno de ellos descrito en un anexo al estándar, introduciéndose otros modos opcionales en la segunda y tercera versión del estándar, H.263+ y H.263++. La aplicación principal de H.263 son las vídeo comunicaciones full-dúplex de baja tasa de bits y de bajo retardo. H.263 puede soportar vídeo comunicaciones a tasas de bits inferiores a 20 Kbps (con una calidad visual muy limitada) y es ampliamente utilizado en videotelefonía y videoconferencia además de en otras muchas nuevas aplicaciones, tales como vídeo basado en Internet.

- H.264

  La continua aparición, un tanto desordenada, de diferentes estándares de compresión de vídeo con el objetivo de mejorar las prestaciones de sus antecesores, así como la excesiva especialización de éstos para un conjunto específico de aplicaciones, llevó al grupo VCEG al comienzo del desarrollo de un nuevo estándar que presentara prestaciones de compresión superiores al resto para todo tipo de aplicaciones, denominándose dicho estándar H.26L, significando la letra `L´, *Long-term Standard*. Una vez comenzado el desarrollo del estándar, la ISO realizó en el año 2001 una llamada abierta a la comunidad científica internacional en búsqueda de nuevas soluciones dentro del campo de la codificación de vídeo avanzada (AVC-*Advanced Video*

Coding). El objetivo de dicho proceso era encontrar nuevas soluciones que permitieran mejorar las prestaciones del estándar MPEG-4 que, aunque prometedor, no había causado un impacto en el mercado similar al de MPEG-2. Entre las propuestas recibidas se encontraban los primeros trabajos realizados por el ITU-T dentro del marco del estándar H.26L, resultando ser la propuesta más innovadora a juicio de los expertos del grupo ISO/IEC. Este hecho provocó que ambas organizaciones se fundieran en un nuevo grupo de trabajo conjunto (JVT-*Joint Video Team*) con el objetivo de crear "una solución única para la próxima generación de estándares de codificación de vídeo". Los esfuerzos del grupo de trabajo JVT, dieron lugar al desarrollo de un nuevo estándar denominado H.264/AVC [6], también conocido como JVT, H.26L o MPEG-4 part 10. Este estándar consigue reducir, por término medio, las tasas de transmisión requeridas por un codificador MPEG-2 en un 50%, y en un 35 % las correspondientes a codificadores basados en los estándares H.263 o MPEG-4.

De los estándares mencionados previamente, los dos estándares desarrollados de manera conjunta han tenido un gran impacto en el mercado y los podemos encontrar en una gran variedad de productos. Entre estos estándares, no cabe duda que el estándar H.264 es el que mejor prestaciones ofrece. El objetivo de este proyecto fue crear un estándar capaz de proporcionar una buena calidad de imagen con tasas de bit sustancialmente menores que los estándares previos sin un incremento excesivo de la complejidad, de forma que su implementación no fuera inviable. Otra de las metas consistía en que el estándar fuera lo suficientemente flexible para ser utilizado en una gran variedad de aplicaciones (altas y bajas tasas de bit, diferentes resoluciones de vídeo) y para trabajar correctamente en una amplia variedad de redes y sistemas (radiodifusión, almacenamiento DVD, redes RTP/IP o sistemas de telefonía multimedia ITU-T). Como se puede observar en la Figura 1.1, tomando como referencia la secuencia de prueba "*Tempete*" el codificador H.264/AVC ofrece mejor calidad de imagen (Y-PSNR), para una determinada tasa de compresión (*bit_rate*), que los estándares MPEG-4, MPEG-2 y H.263. El estándar MPEG-2 se utiliza especialmente en difusión de vídeo, tanto vía satélite como en formato DVD. El estándar MPEG-4 mejora las prestaciones del MPEG-2, añadiendo funciones tales como la estimación de movimiento global o la posibilidad de segmentar la imagen en objetos visuales, posibilitando un tratamiento personalizado a cada objeto. Este estándar es más

conocido a nivel de usuario como DivX o Digital Vídeo Express. Finalmente, H.263 es un estándar usado para videoconferencia de muy bajo ancho de banda.

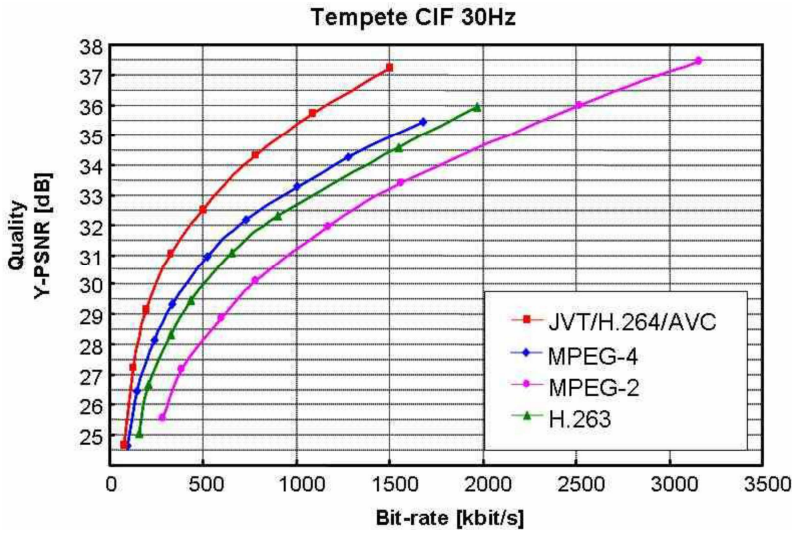

Figura 1 Comparación de las prestaciones del estándar H.264/AVC con los estándares de compresión MPEG-4, MPEG-2 y H.263 codificando la secuencia "Tempete" a 30 fotogramas por segundo

El H.264 impulsó un gran salto en el área de vídeo, muchos sustituyeron sus códecs de compresión propios por éste. Se emplea en productos y servicios de empresas como Adobe, Apple, BBC, BT, France Telecom, Intel, Microsoft, Motorola, Nokia, Polycom, Samsung, Sony, Toshiba y otras, para transmitir vídeo de alta definición por televisión de radiodifusión, TV por cable y por diversos servicios de televisión directa por satélite, así como en discos Blu-Ray, telefonía móvil, videoconferencias, medios de almacenamiento digital y televisión por Protocolo Internet (TVIP). En definitiva, es el estándar de codificación de vídeo más empleado en el mundo actualmente. Pero, ¿es el H.264 lo suficientemente bueno para la carga presente en las redes mundiales, donde, según estudios, los vídeos consumen más de la mitad del ancho de banda? En respuesta a esta pregunta nace el nuevo estándar de codificación de vídeo HEVC.

¿Qué es HEVC? High Efficiency Video Coding también conocido como H.265 es el nuevo estándar de compresión de vídeo desarrollado por el grupo Joint Collaborative Team on Video Coding (JCT_VC). Este grupo está formado por los grupos ISO/IEC Moving Pictures Experts (MPEG) e ITU-T Video Coding Experts (VCEG), como ya se comentó en párrafos anteriores, estos dos grupos fueron los mismos que se unieron para desarrollar de manera conjunta el estándar predecesor, el estándar H.264. Este nuevo estándar aliviará considerablemente la

carga en las redes mundiales dado que, según las estimaciones, la transmisión de vídeo representa la mitad del ancho de banda empleado por su antecesor, H.264, para mantener la misma calidad. H.264 representa actualmente más del 80% de todo el vídeo en la web. La nueva norma desencadenará una nueva fase de innovación en la producción de vídeo en toda la gama de las TIC, desde los dispositivos móviles hasta la TV de ultra alta definición. Empresas como ATEME, Broadcom, Cyberlink, Ericsson, Fraunhofer HHI, Mitsubishi, NHK, NTT DOCOMO y Qualcomm ya han presentado modelos que emplean HEVC.

Tomando como referencia la secuencia de prueba "*Johnny*", en la Figura 2 se observa que el codificador HEVC ofrece la misma calidad de imagen (Y-PSNR) usando la mitad de tasa de compresión (*bit_rate*) que el estándar H.264.

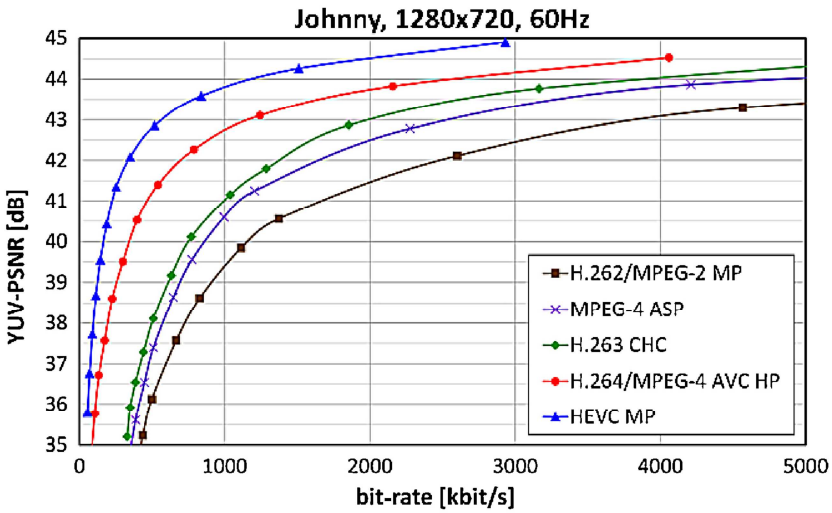

Figura 2 Comparación de las prestaciones del estándar HEVC con estándares previos codificando la secuencia "Johnny" a 60 fotogramas por segundo

Debido al impacto que puede tener este nuevo estándar en el mundo del vídeo, se ha considerado de notable interés su estudio y caracterización.

1.3. Objetivos

Los objetivos principales de este Trabajo Fin de Máster se presentaron en la memoria del anteproyecto del mismo y se vuelven a desarrollar de manera extendida a continuación.

- Estudiar el estándar HEVC. Entender la meta de este nuevo estándar y qué nuevos cambios se han introducido para alcanzar el propósito deseado que no es nada menos que alcanzar la misma calidad de vídeo que su predecesor, H.264, usando la mitad de tasa de bit (bit rate).
- Comparativa con H.264 en términos de arquitectura planteada.
- Estudio del modelo software de referencia y sus parámetros.
- Familiarización con el software y los parámetros del modelo para posteriormente crear los bancos de test que se usarán para caracterizar el estándar.
- Elección de las secuencias de imágenes que se van a emplear en los bancos de test. Estos vídeos se elegirán de manera que tengamos un amplio abanico de características, por ejemplo, frames con mucho movimiento, frames con fodo estático y movimiento en el primer plano, etc.
- Pruebas y análisis de los resultados obtenidos. Las pruebas se realizarán empleando diferentes QPs y condiciones de test. De este modo se podrá observar el comportamiento del codificador frente a diferentes configuraciones. Para el análisis de los resultados emplearemos medidas de PSNR frente a bit rate, además de evaluar qué etapas muestran mayor complejidad desde el punto de vista de tiempo de ejecución.
- Conclusión. Por último, se desarrollarán las conclusiones teniendo en cuenta tanto el estudio teórico del estándar HEVC como las pruebas realizadas empleando el modelo software de referencia.

Para conseguir los objetivos aquí presentados, se empleará la recomendación presentada por la ITU: High Efficiency Video Coding [7] y el código software de referencia HM 8, ambos se pueden obtener a través de la web del instituto de Fraunhofer (http://hevc.hhi.fraunhofer.de). La caracterización del estándar se realizará mediante curvas PSNR (calidad) frente a bit rate (tasa de bit).

1.4. Organización de la memoria

La memoria está estructurada en cinco capítulos y la bibliografía.

En el primer capítulo se realizará la introducción al Trabajo Fin de Máster (TFM). Se presenta una breve reseña sobre la historia de los estándares de compresión de vídeo y la razón de desarrollar este TFM. Posteriormente, se procederá a enumerar los objetivos propuestos y, finalmente, se presenta la estructura de este documento.

En el segundo capítulo se presenta el estudio del estándar HEVC, principales bloques, similitudes y diferencias con respecto a los estándares anteriores y perfiles disponibles.

En el tercer capítulo se presenta el estudio del software de referencia así como los parámetros principales con los que vamos a trabajar en los bancos de test.

En el cuarto capítulo se presentan los bancos de test y las secuencias que se van a emplear para la caracterización del estándar.

En el quinto capítulo se presentarán los resultados obtenidos. Finalmente, se hace un resumen de los objetivos cumplidos y se presentan las conclusiones finales del TFM.

Las referencias documentales empleadas en este documento se detallan en la bibliografía.

Capítulo 2: Estándar de compresión de vídeo HEVC

Todos coinciden en que el nuevo estándar HEVC/H.265 es el códec del futuro, pero una imagen poco clara y la escasez de dispositivos compatibles dan a entender que existe un largo camino por recorrer antes de la adopción de este estándar. H.265/High Efficiency Video Coding (HEVC) es el sucesor del códec H.264. Ambos han sido desarrollados de manera conjunta por ISO/IEC Moving Picture Experts (MPEG) e ITU-T Video Coding Experts Group (VCEG). La principal meta de este nuevo estándar de codificación de vídeo es alcanzar una eficiencia de compresión un 50% mejor que la de su predecesor, H.264.

2.1. Estandarización

El proceso de estandarización del estándar HEVC fue un proceso de 9 años. Después de la finalización del H.264/MPEG-4 AVC High Profile a mediados de 2004, ITU-T VCEG comenzó a trabajar para identificar cuáles serían los posibles futuros avances. En octubre de ese mismo año, 2004, presentaron varias técnicas que podrían suponer una mejora en base a H.264. Al igual que hicieron para H.264, empezaron a trabajar en un software de referencia del cual pudieran ir extrayendo resultados. A comienzos del 2005, empezaron a evaluar los resultados de lo que llamaron KTA (Key Technical Areas) codebase. Este software está basado en el software de referencia desarrollado para H.264. Entre 2004 y 2008 este software fue evolucionando, se fueron integrando nuevas propuestas, al mismo tiempo que se iban analizando los resultados obtenidos. La idea de ITU-T era desarrollar un estándar que fuera capaz de reducir a la mitad la tasa de bit conservando la misma calidad de imagen que H.264.

En 2007, el grupo ISO/IEC MPEG comenzó un proceso similar al de ITU-T VCEG. Basados en el software KTA de VCEG, en julio de 2009 obtuvieron resultados de reducción de la tasa de bit de un 20% comparado con H.264/AVC High Profile. Estos resultados hicieron que los dos grupos consideraran nuevamente colaborar para desarrollar y estandarizar un nuevo estándar de vídeo.

En 2010, VCEG y MPEG organizaron un evento llamado “Call for Proposals”. Para este recibieron varias propuestas, algunas de ellas demostraban que la reducción de tasa de bit, manteniendo la misma calidad de imagen, era realizable.

Además de obtener interesantes propuestas, en este evento se adoptó el nombre High Efficiency Video Coding (HEVC) para el nuevo estándar. A finales de 2010, octubre, se presentó el primer borrador de trabajo de dicho estándar. Después de 2 años, el febrero de 2012, y basándose en el sexto borrador se aprobó éste como el primer borrador oficial por el Committee Draft de HEVC, el último aprobado y que fue considerado como una posible primera versión del estándar fue el octavo draft. En junio de 2012, comenzó el proceso de patentado.

En abril de 2013, HEVC/H.265 [7] fue aprobado por la ITU-T y en junio fue formalmente publicado. En esencia, esto significa que la primera versión del estándar es congelada y, por tanto, varias empresas desarrolladoras podrán finalizar sus primeros productos para HEVC.

En Tabla 1 se resumen las principales fechas en el proceso de estandarización.

Fecha	Evento
2004	VCEG comienza a trabajar para desarrollar un nuevo estándar
2007	MPEG comienza a trabajar con el mismo objetivo que VCEG
2009	VCEG y MPEG deciden trabajar de manera conjunta formando el grupo Joint Collaborative Team on Video Coding (JCT-VC)
2010	Toma lugar el evento “Call for Proposals” del cual se extraen buenas propuestas y del cual nace el nombre HEVC
Feb. 2010	Se acepta el primer draft (Committee Draft)
Jul. 2012	Se acepta el último draft
Abr. 2013	HEVC/H.265 es aprobado
Jun. 2013	HEVC/H.265 es publicado

Tabla 1 Hitos importantes en el proceso de estandarización

2.2. Tecnología y características clave

HEVC emplea esencialmente la misma estructura híbrida (predicción espacial/temporal y 2-D transform coding) usada en todos los estándares de compresión de vídeo desde H.261 pero con mejoras en cada una de las etapas. En la Figura 3, se puede observar el diagrama de bloques para un codificador HEVC y como se basa en la ya nombrada estructura híbrida.

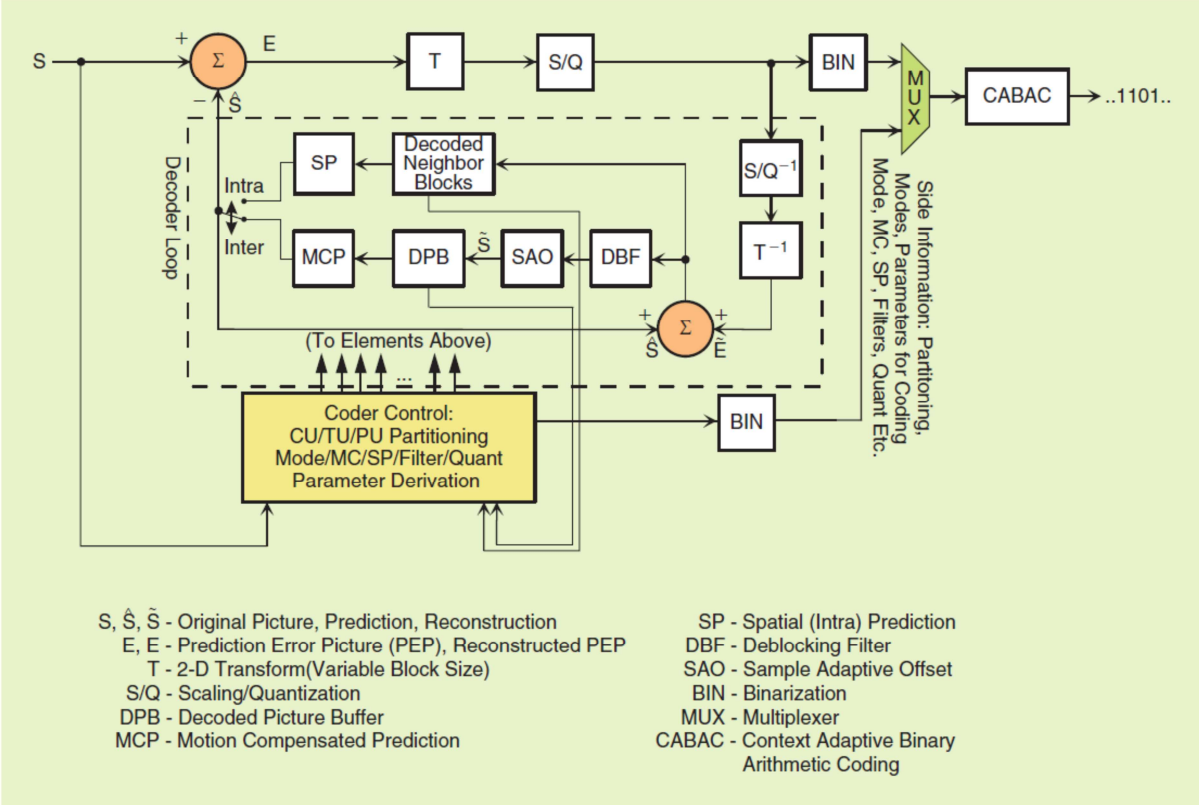

Figura 3 Codificador de vídeo híbrido para HEVC [8]

El flujo producido por un codificador HEVC sería, a grandes rasgos, el siguiente: cada imagen es dividida en bloques (más adelante se detallará cómo se puede realizar esta división). Para cada una de estas divisiones, se predice una imagen en base a la información previamente decodificada, esta información puede pertenecer a imágenes previamente decodificadas (predicción temporal o inter) o a muestras de la misma imagen previamente decodificadas (predicción espacial o intra). La primera imagen de una secuencia de vídeo se codifica usando la predicción intra. Para el resto de las imágenes se puede usar intra o inter, siendo la predicción temporal (inter) la más usada en estos casos. A continuación, el residuo (E) de la predicción espacial o temporal, el cual es la diferencia entre la

imagen original y la predicha, es transformado (T) y los coeficientes transformados son posteriormente escalados, cuantificados (S/Q), codificados entrópicamente (CABAC) y transmitidos como parte del bit stream. Los vectores de movimiento y los parámetros adicionales que también tienen que ser transmitidos para llevar a cabo la decodificación son también codificados y transmitidos como parte del bit stream. Como se puede observar en la figura, el codificador duplica el bucle de decodificación para poder realizar las predicciones.

A priori puede parecer que HEVC es bastante similar al resto de estándares previos pero existen diferencias claves que permiten que este estándar haya mejorado la capacidad de compresión de vídeo, estas características claves van a ser explicadas en los siguientes sub-apartados.

2.2.1. Coding Tree Units y Coding Tree Block Structure

En estándares anteriores, la unidad de codificación era el macrobloque. Un macrobloque contiene un bloque 16x16 de luma y, en el caso de configuración 4:2:0, dos bloques 8x8 para las muestras de croma. Por el contrario, HEVC usa Coding Tree Unit (CTU), esta estructura sería la análoga al macrobloque. La CTU es la unidad básica de procesamiento y consiste en Coding Tree Blocks (CTBs) para luma y croma. Un CTB de luma cubre LxL muestras de luma y las correspondientes CTBs de croma L/2xL/2 muestra de cada una de las dos componentes de croma. El valor de L puede ser 16, 32 o 64, con tamaños más grandes obtenemos más compresión. El macrobloque usado en H.264 era de tamaño fijo, 16x16, HEVC soporta tamaños variables que pueden ser seleccionados de acuerdo a las necesidades del codificador en términos de memoria y coste computacional, ofreciendo así mayor flexibilidad, tal y como se puede ver en la Figura 4.

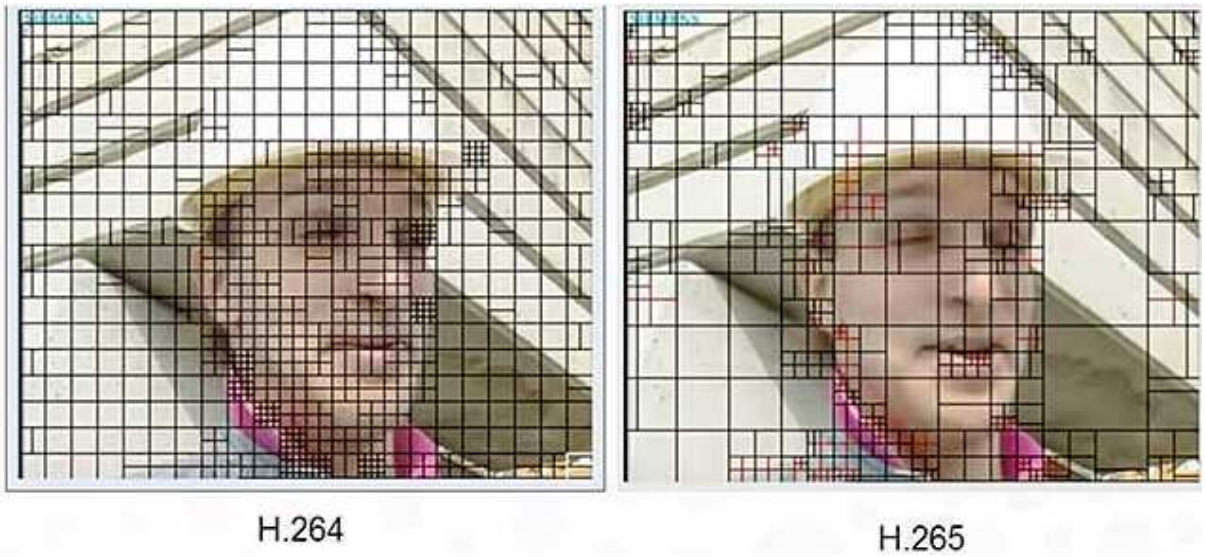

Figura 4 Particionado en H.264 (izquierda) y en HEVC/H.265 (derecha)

Los CTBs de luma y croma pueden ser directamente usados como Coding Blocks (CBs) o puede ser particionados nuevamente formando distintos CBs (siempre cuadrados y tamaño mínimo de 8x8). Este particionado usa estructuras en árbol (tree structure), tal y como se verá posteriormente en la Figura 6. En general, esta partición es aplicada de igual manera a las muestras de luma y croma pero existen algunas excepciones cuando se alcanzan algunos valores mínimos para croma. Un CB de luma junto con los CBs de croma y los elementos de sintaxis asociados es a lo que nos referiremos como Coding Unit (CU).

A nivel de CU se pueden realizar divisiones adicionales en Prediction Units (PU) y Transform Units (TU). A nivel de CU se decide si la predicción va a ser espacial (intra) o temporal (inter). Si el modo de predicción es intra, el tamaño del Prediction Block (PB) es el mismo que el de CB excepto para el tamaño más pequeño de CB, en este caso, el CB puede ser dividido en cuatro PB (Prediction Blocks) y, por tanto, tener cada uno de ellos un mode de predicción intra diferente. La razón de tener esta excepción es permitir tener distintos modos de predicción intra para los bloques de tamaño 4x4. Cuando los bloques de luma son 4x4, los bloques de croma también son 4x4. Si el modo de predicción es determinado como temporal (inter), los CBs de luma y croma pueden ser divididos en 2 o 4 PBs. La división en 4 es sólo permitida si el tamaño del CB es igual al mínimo permitido. En la Figura 5, podemos observar las diferentes posibilidades de dividir un CB en PBs en este caso.

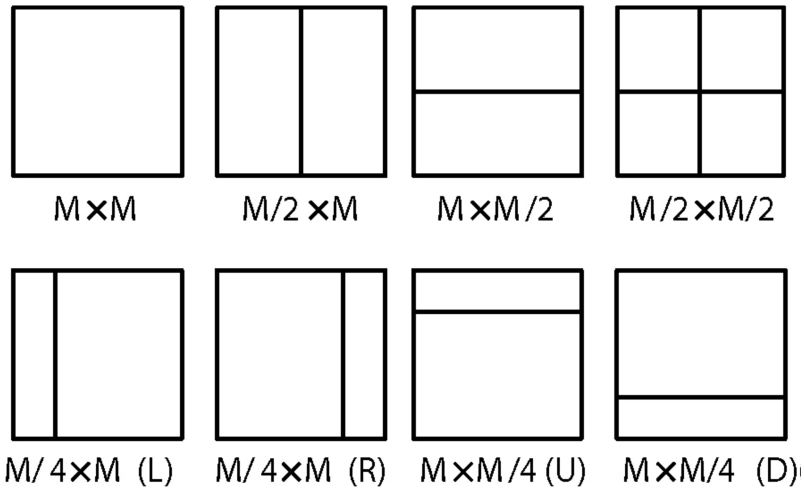

Figura 5 Modos de división de un CB en PBs (sujeto a restricciones). Para intra, sólo son soportadas las divisiones MxM y M/2xM/2

La parte de debajo de la Figura 5 muestra los llamados Asymmetric Motion Partitioning (AMP). Estas divisiones son sólo permitidas cuando M es mayor o igual a 16 para la luma.

En predicción inter, cada PB tiene asignado uno o dos vectores de movimiento e índices que apuntan a la imagen de referencia. Para aliviar memoria, PBs de luma con tamaño 4x4 no son permitidos en predicción inter.

Los PBs de luma y croma, junto con la sintaxis de predicción asociada forman la Prediction Unit (PU).

2.2.2. Transform Units y Transform Blocks

Como ya se ha apuntado anteriormente, el residuo resultante de restar la imagen original con la predicha es transformado, escalado, cuantificado y codificado. Para ello, un CB puede ser particionado recursivamente en Transform Block (TBs). Al igual que los CBs, los TBs sólo pueden ser divididos en cuadrados, tal y como se observa en la Figura 6.

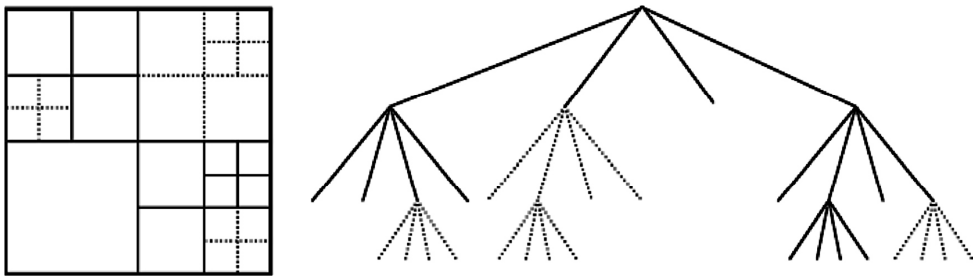

Figura 6 División de un CTB en CBs (líneas sólidas) y TBs (líneas discontinuas) (izquierda) y su correspondiente quadtree (derecha)

Para un CB de luma MxM, se indica, mediante el uso de un flag, si va a ser dividido en cuatro bloques M/2xM/2. Si la división puede ser aún mayor, tal y como se debe indicar en el parámetro que corresponda a máxima profundidad, cada cuadrante tendrá un flag que indique si debe ser también dividido en cuatro cuadrantes. Los nodos que formen las hojas del árbol serán los que serán finalmente procesados.

A diferencia con previos estándares, HEVC permite que un TB abarque múltiples PBs en una CU inter. De este modo se maximizan los beneficios de la codificación.

2.2.3. Predicción Espacial (Intra)

La predicción espacial en el estándar HEVC opera de manera similar a H.264/AVC pero extendiendo el número de posibles direcciones a usar de manera significativa. HEVC usa 33 modos de predicción con diferentes direcciones comparado con los 8 modos definidos en H.264. Este modo es denominado como Intra_Angular[k], donde k toma valores desde 2 a 34. Los ángulos están diseñados intencionadamente para proporcionar una cobertura más densa para los ángulos cercanos a la horizontal y vertical y una cobertura más gruesa para los ángulos más cercanos a la diagonal. Cuando se usa este modo, la predicción se hace para TBs de tamaño 4x4 hasta 32x32, siempre cuadrados. Las muestras de los vecinos Transform Blocks (TBs) reconstruidos (pero no aún filtradas por los filtros in-loop) son usados como datos de referencia para la predicción. En la Figura 7 se muestran las diferentes orientaciones para el modo direccional.

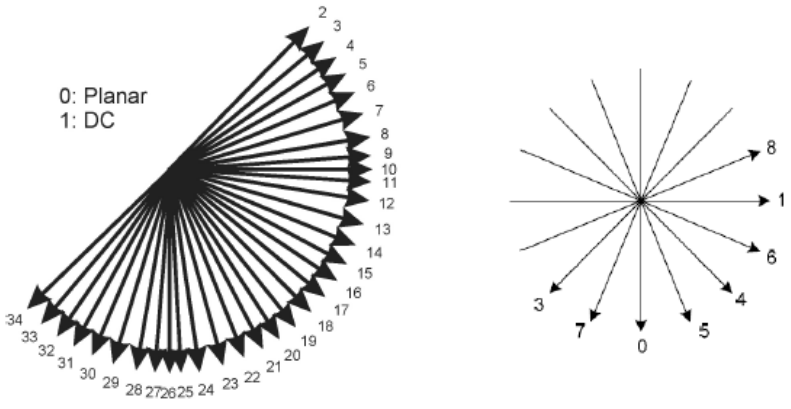

Figura 7 Modos y orientaciones direccionales empleados en la predicción espacial para HEVC (izquierda) y H.264 (derecha)

Al igual que este último, HEVC también define un modo de predicción plana (montaje superficial, donde se asume una superficie con inclinación horizontal y vertical derivada en base a las muestras vecinas), denominado Intra_Planar, y uno de predicción DC (asumiendo una superficie plana con valor igual al de la media de las muestras vecinas), denominado Intra_DC.

Para el caso de la croma, los modos horizontal, vertical, plano y DC pueden ser explícitamente indicados para ser usados o también se puede indicar que siga el mismo modo empleado para la luma. En casos especiales, para evitar redundancia, la croma puede emplear el modo izquierda-hacia abajo diagonal.

2.2.4. Predicción Temporal (Inter)

Teniendo en cuenta la partición empleada, HEVC soporta más formas de Prediction Blocks (PBs) que sus antecesores. Los modos de partición PART_2Nx2N, PART_2NxN y PART_Nx2N indican los casos cuando el Coding Block (CB) no es particionado, cuando es dividido en dos partes iguales horizontalmente y cuando es dividido en dos partes iguales verticalmente, respectivamente. PART_NxN indica que el CB es dividido en cuatro PBs de igual tamaño, este modo es sólo soportado cuando el tamaño del CB es el mínimo permitido. También existe la posibilidad de dividir el CB de manera asimétrica. Para más detalle, se remite a la sección 2.2.1.

Al igual que H.264/AVC, la compensación de movimiento para la predicción temporal en HEVC usa precisión de cuarto de pixel para la luma. Para las muestras de croma, la precisión del vector de movimiento viene determinada por el formato empleado, que en el caso de emplear 4:2:0 sería de un octavo de píxel. Para la interpolación de las muestras de luma, HEVC usa un filtrado de 8 etapas para obtener las muestras de ½ píxel y un filtrado de 7 etapas para obtener las de ¼ de píxel. En contra, H.264 usa filtrado de 6 etapas para obtener las muestras de medio pixel seguido de una media entre muestras de medio pixel para obtener las de cuarto de pixel. En la Figura8, las posiciones en mayúsculas (Ai,j) representan las muestras de luma disponibles en posiciones enteras de píxel, mientras que las posiciones en minúsculas representan las muestras en posiciones no enteras, las cuales hay que generar usando interpolación. En la Tabla 2 se muestran los coeficientes de los filtros empleados para la generación de las muestras de medio y cuatro de píxel.

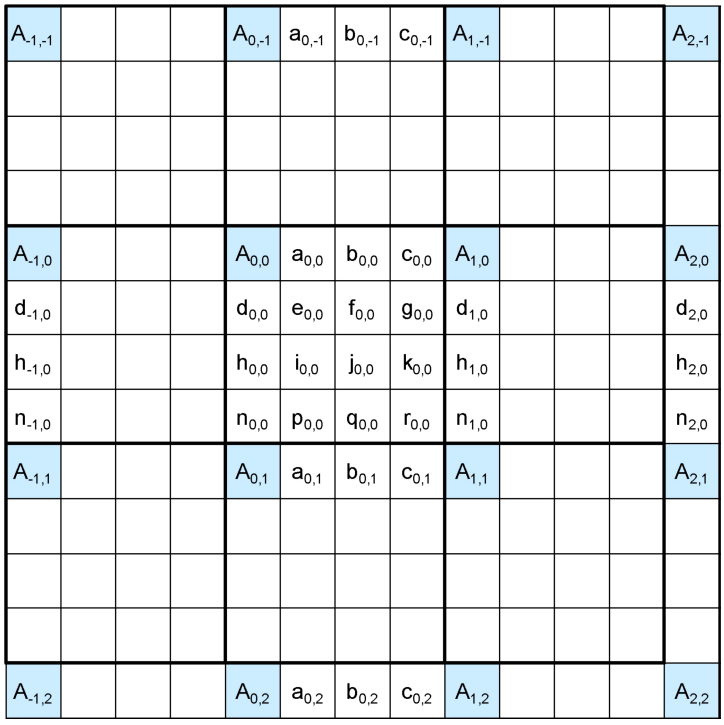

Figura 8 Muestras en posiciones enteras y fraccionales para la interpolación de la luma

Índice i	-3	-2	-1	0	1	2	3	4
hfilter[i]	-1	4	-11	40	40	-11	4	1
qfilter[i]	-1	4	-10	58	17	-5	1	

Tabla 2 Coeficientes empleados para los filtros de 8 y 7 etapas para la luma, medio píxel (hfilter) y cuarto de píxel (qfilter), respectivamente

De manera similar a H.264/AVC, se usan varias imágenes de referencia, hasta 6 imágenes (si se emplean imágenes de menor resolución este número puede incrementarse hasta 16). Por cada Prediction Block (PB), uno o dos vectores de movimiento pueden ser transmitidos, resultando de esta manera en codificación uni-predictiva o bi-predictiva, respectivamente. En el caso de emplear dos vectores de movimiento, el resultado se puede obtener con la media de las predicciones o poniendo un peso a cada una de ellas.

En el caso de HEVC, es usado el modo de predicción avanzado (Advantage Motion Vector), aumentando el número de posibles candidatos basándose en datos desde los PBs adyacentes y la región donde se encuentra en la imagen de referencia. Se define también un nuevo modo de mezclado (merge mode), indicando la herencia de vectores de movimiento desde los PBs vecinos en el mismo CTU o imagen de referencia.

Para las muestras de croma, el proceso de interpolación de muestras no enteras, fraccionales, es similar al usado para la componente de luma, pero empleando 4 etapas. HEVC define un set de filtros de 4 etapas para las muestras de octavo de píxel (formato 4:2:0) tal y como se observa en la Tabla 3. En contra, H.264/AVC, emplea filtrado de 2 etapas bilineal.

Índice	-1	0	1	2
filter1[i]	-2	58	10	-2
filter2[i]	-4	54	16	-2
filter3[i]	-6	46	28	-4
filter4[i]	-4	36	36	-4

Tabla 3 Coeficientes empleados para los filtros de 4 etapas para las muestras fraccionales de croma

Continuando con la comparación con H.264/AVC, se debe apuntar que los modos directo y skip son mejorados.

2.2.5. Codificación entrópica

HEVC usa un único método de codificación entrópica, CABAC (Context Adaptive Binary Arithmetic Coding), en lugar de dos posibles codificadores como en H.264/AVC, CAVLC (Context Adaptive Variable Length Coding) o CABAC dependiendo del perfil empleado. El núcleo empleado en HEVC es similar al empleado en H.264/AVC pero ha sufrido algunas modificaciones usando un esquema de selección de contexto más sofisticada, la cual es particularmente eficiente para los TBs de mayor tamaño. El objetivo de estas modificaciones es mejorar la velocidad del throughput, la eficiencia de la compresión y los requisitos de memoria. A continuación se presenta puntos característicos de como HEVC hace uso de CABAC.

- Modelado del contexto: Un factor clave para mejorar la eficiencia de la codificación CABAC es seleccionar adecuadamente el contexto. En HEVC, el uso de las divisiones en coding tree y en transform tree se explota para obtener más índices de los que se obtenían en H.264/AVC. Y a pesar de que el número de contextos empleados en HEVC es menor, la compresión obtenida es mejor que en H.264/AVC. Además, en HEVC se hace mayor uso del modo bypass para incrementar el throughput ya que se reduce el número de datos que necesita ser codificado usando CABAC.

- Búsqueda de coeficientes adaptativos: la búsqueda de estos coeficientes se realiza en sub-bloques 4x4 para todos los TBs. Los métodos de búsqueda empleados son 3, diagonal up-right, horizontal y vertical, tal y como se puede ver en la Figura 9. La selección del método de escaneo depende de la dirección de la predicción espacial (intra). Si la dirección es próxima a ser horizontal se emplea el scan vertical, si es próxima a ser vertical se emplea el scan horizontal, para el resto de casos se emplea el diagonal up-right.

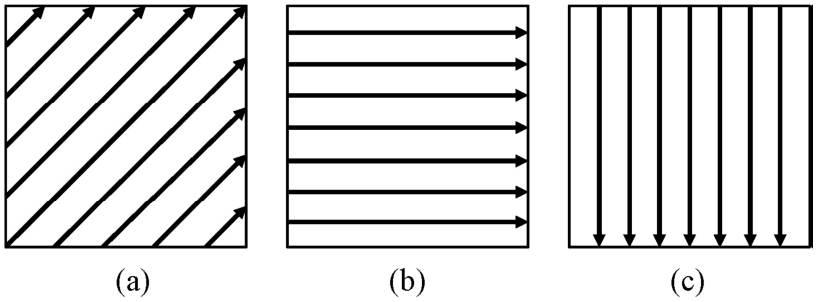

Figura 9 Métodos de escaneo en HEVC. (a) Diagonal up-right. (b) horizontal. (c) vertical

2.2.6. Filtros In-Loop

En HEVC, se han integrado dos filtros, tal y como se puede apreciar en la Figura 3 (apartado 2.2). Un deblocking filter (DBF) seguido de un filtro Sample Adaptive Offset (SAO) son aplicados para reconstruir las muestras antes de que sean almacenadas en el buffer de muestras decodificadas (DPB) en el bucle de decodificación.

El DBF reduce el efecto de bloque que se genera al emplear la codificación basada en bloques. El DBF usado en HEVC es similar al empleado por H.264/AVC, pero los procesos de toma de decisión y filtrado han sido modificados para favorecer procesos en paralelo. El número de bordes de luma a ser filtrado en el peor de los casos ha sido reducido a la mitad, esto se ha logrado limitando las operaciones del DBF a operar con un grid en los bordes de 8x8, en lugar de 4x4.

El filtro SAO, por el contrario, es un elemento completamente nuevo añadido en HEVC. Mientras que el DBF es únicamente aplicado a las muestras localizadas en los bordes de los bloques, el filtro SAO es aplicado de manera adaptativa a todas las muestras. El objetivo de este filtro es mejorar la reconstrucción de las muestras decodificadas sumando un valor de offset a cada muestra después de la aplicación del DBF. Estos offsets dependen de valores transmitidos por el codificador y almacenados en tablas (lookup tables - LUTs). Este mapeado no es lineal. Este filtrado tiene dos opciones de uso e incluso se puede desactivar, esto se puede hacer por Coding Tree Block (CTB) usando el elemento de sintaxis sao_type_idx; donde 0, indica que este filtrado no es

aplicable; 1, indica que se usará SAO con el modo denominado “band offset”; y 2, SAO “edge offset”. Usando “band offset” (sao_type_idx=1), el valor de offset seleccionado depende directamente del valor de amplitud de la muestra. En este modo, se reduce el efecto generado artificialmente en los contornos en la decodificación. En el caso de “edge offset” (sao_type_idx=2), se tiene que también proporcionar sao_eo_class con valores entre 0 y 3, estos valores indican la clasificación del offset usado para el CTB, pudiendo ser horizontal, vertical o uno o dos direcciones de gradiente diagonal, respectivamente. En la Figura 10 se pueden observar los 4 gradientes empleados. Con “p” se indica el centro de la muestra a ser considerado y con “n0” y “n1” dos muestras vecinas. Cada muestra en el CTB es clasificada en uno de las cinco EdgeIdx categorías. Para ello, se compara el valor de la muestra p con las muestras vecinas n0 y n1 tal y como se indica en la Tabla 4. Dependiendo de estos valores se extrae un offset de la lookup table y se suma a la muestra. Para las categorías 1 y 2, los valores son siempre positivos, mientras que para las categorías 3 y 4 son siempre negativos. De esta manera, el filtro suaviza los efectos en los bordes. En definitiva, SAO es un filtro no lineal que permite un refinamiento adicional al ya proporcionado por el DBF.

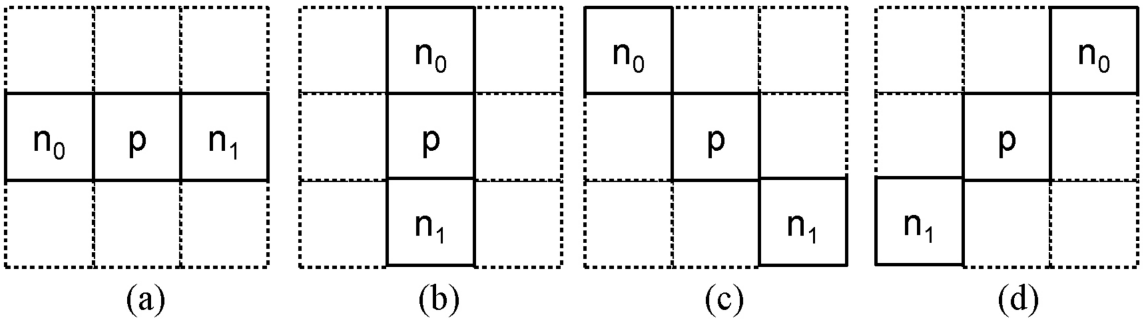

Figura 10 Cuatro gradientes usados en SAO modo “edge offset”. (a) horizontal. (b) vertical. (c)135º diagonal . (d) 45º diagonal

EdgeIdx	Condición	Significado
0	Casos no indicados abajo	Área monótona
1	p<n0 y p<n1	Mínimo local
2	p<n0 y p=n1 ó p<n1 y p=n0	Borde
3	p>n0 y p=n1 ó p>n1 y p=n0	Borde
4	p>n0 y p>n1	Máximo local

Tabla 4 EdgeIdx categorías para SAO en modo “edge offset”

Durante el desarrollo del estándar HEVC, se barajó la posibilidad un tercer filtro llamado ALF (Adaptive Loop Filter) justo después del SAO, pero en el diseño final del codificador se decidió eliminar.

2.2.7. Perfiles

Al igual que en estándares previos, los conceptos de "perfiles" y "niveles" son usados en HEVC, los cuales son diseñados para facilitar la inter-operatividad entre varias aplicaciones del estándar que tienen los mismos requisitos de funcionalidad. Un perfil define un conjunto de herramientas de codificación o algoritmos que pueden ser usados en la generación de un determinado bit stream, mientras que un nivel fija una serie de restricciones a través de parámetros en el bit stream que corresponden con la carga del proceso de decodificación y la memoria. Actualmente, tres perfiles han sido definidos en la primera versión del estándar HEVC: Main, Main 10 y Main Still Picture, éstos serán los que explicaremos en más detalle pero cabe recalcar que en agosto se presentó un borrador con cinco perfiles adicionales para futuras versiones. En la Tabla 5 se resumen los perfiles del estándar HEVC.

Feature	Version 1		Range extensions				
	Main	Main 10	Main 12	Main 4:2:2 10	Main 4:2:2 12	Main 4:4:4 10	Main 4:4:4 12
Bith depth	8	10	8-12	8-10	8-12	8-10	8-12
Chroma sampling formats	4:2:0	4:2:0	4:2:0	4:2:0 4:2:2	4:2:0 4:2:2	4:2:0 4:2:2 4:4:4	4:2:0 4:2:2 4:4:4
4:0:0 (monochrome)	No	No	Yes	Yes	Yes	Yes	Yes
Intra block copy	No	No	Yes	Yes	Yes	Yes	Yes
Intra smoothing disabling	No	No	Yes	Yes	Yes	Yes	Yes
Residual DPCM inter/intra	No	No	Yes	Yes	Yes	Yes	Yes
Transform skip block size larger than 4x4	No	No	Yes	Yes	Yes	Yes	Yes
Transform skip context/rotation	No	No	Yes	Yes	Yes	Yes	Yes
MinCR reduced to half its based value	No	No	No	Yes	Yes	Yes	Yes
Extended precision processing	No	No	No	No	No	No	No
Separate color plane	No	No	No	No	No	No	No

Tabla 5 Perfiles para HEVC

El perfil Main es el empleado para típicas aplicaciones de vídeo, permite una profundidad de 8 bits por bit con formato 4:2:0. Este es el formato más comúnmente empleado en los dispositivos de consumo.

El perfil Main 10 permite una profundidad de bit desde 8 a 10 bits. Es empleado para vídeos donde necesitamos mayor profundidad de bit (10 bits). Una mayor profundidad de bit permite un mayor número de colores. Un mayor número de colores permite a su vez tener transiciones más suaves entre distintos colores. En definitiva, este perfil permite alcanzar una calidad de vídeo mayor que la que proporcionada por el perfil Main. Un numeroso grupo de empresa apoya la introducción de este perfil ya que tal y como apunta Ericsson e Imagination Technologies serán necesarios 10 bits para dar soporte a la televisión de ultra alta definición (UHDTV). Entre estas empresas se encuentran ATEME, BBC, CISCO, Ericsson, Motorola Mobility, Thomson Video Networks, etc.

Y, por último, el perfil Main Still Picture es empleado para imágenes (con predicción intra únicamente). Con este perfil se codifica con las mismas restricciones que el perfil Main, es un subset de este último. En abril de 2012, se realizó un estudio para comparar los resultados obtenidos con el perfil Main Still Picture y se observó que con HEVC Main Still Picture se conseguía una reducción de la tasa de bit significativa comparado con JPEG, obsérvense resultados en la Tabla 6.

Estándar de codificación (método de test)	Reducción media de la tasa de bit comparado con	
	JPEG 2000	JPEG
HEVC (PSNR)	20,26%	61,63%
HEVC (MOS)	30,96%	43,10%

Tabla 6 Comparativa entre HEVC Main Still Picture y JPEG 2000 y JPEG en cuanto a tasa de bit basándose en PSNR (Peak Signal to Noise Ratio) y MOS (Mean Of Score)

Capítulo 3: Código de referencia

Fraunhofer HHI es una organización orientada al desarrollo de aplicaciones digitales. Su labor se concentra en diseños orientados a la tecnología de la información, la mayoría de ellos nuevos productos que desarrollan conjuntamente con industrias mediante colaboraciones. Una de sus áreas principales es la codificación de vídeo y, tal y como hicieron para el estándar H.264/AVC, el departamento de procesado de imagen de este instituto ha desarrollado una solución software para demostrar los beneficios del nuevo estándar de vídeo HEVC. Este software es considerado en el sector como el "Golden reference" del nuevo estándar.

3.1. Introducción

Para el desarrollo de este Trabajo Fin de Máster, se ha tomado como referencia el codificador/decodificar HEVC implementado en software empleado lenguaje C++ por el grupo de desarrollo de procesado de imagen del instituto de Fraunhofer [9]. El propósito principal de esta solución software es el de constituir una herramienta útil con la que explorar las diferentes posibilidades de este nuevo estándar y ser capaces de analizar su comportamiento. Para ello, se ha procedido a estudiar dicho código identificando las principales funciones y parámetros. Estos parámetros serán empleados posteriormente en los bancos de test para construir diferentes escenarios. Los resultados obtenidos de estas pruebas nos permitirán caracterizar el estándar.

El estudio de los parámetros ha sido desarrollado centrándose en el proceso de codificación ya que es en el que se basará la caracterización objetivo de este trabajo.

La versión de código de referencia empleada es la versión HM 8.0 [10]. El software fue descargado de los servidores SVN del instituto. En previas versiones de HM, el codificador contemplaba dos perfiles, *Low Complexity* (LC) y *High Efficiency* (HE). HM 5.0 introdujo el modo de operación HE para 10 bits además de los anteriores HE y LC. A medida que se fue avanzando en el desarrollo del estándar, el código referencia se fue adaptando hasta llegar al HM 8.0. Esta versión no es la última pero contiene una versión estable de los perfiles Main y Main10.

3.2. Módulos principales

El software de referencia HM 8.0 es un código C++ dividido en 6 grandes módulos y en numerosos bloques. Es un código complicado de seguir y para evitar que esta sección se convierta en tediosa, se ha decidido resumir la estructura en las Tablas 7, 8 y 9. La Tabla 7 representa los principales módulos, la Tabla 8 las principales clases y la Tabla 9 los principales ficheros. Es importante resaltar que el código contiene más clases y ficheros pero aquí se detallan los que se han considerado de especial interés para el estudio a realizar en este Trabajo Fin de Máster.

Módulo	Descripción
TLibCommon	Librería que contiene funcionalidad que es usada tanto en el proceso de codificación como en el de decodificación
TLibEncoder	Librería que contiene funcionalidad exclusiva para el proceso de codificación
TLibDecoder	Librería que contiene funcionalidad exclusiva para el proceso de decodificación
TAppCommon	Aplicación para dar soporte a la librería de elementos comunes
TAppEncoder	Aplicación para el proceso de codificación
TAppDecoder	Aplicación para el proceso de codificación

Tabla 7 Módulos del HM 8.0

Clase	Descripción
TDecTop	Decodificador
TEncTop	Codificador
TDecCu	Decodificador para una CU
TEncCu	Codificador para una CU
TEncCfg	Codificación para el codificador
TComTrQuant	Transformación and cuantificación
TComPrediction	Predicción
TComInterpolationFilter	Filtro de interpolación
TComMv	Vectores de movimiento
TComRdCost	Coste de la tasa de distorsión (RD – Rate Distortion)
TComRdCostWeighttPrediction	Coste de la tasa de distorsión, con predicción con peso
TComWeightPrediction	Predicción con pesos
TComCABACTables	Contiene las tablas empleadas por CABAC
TEncBinCABAC/TDecCABAC	CABAC
TDecEntropy	Decodificador entrópico
TEncEntropy	Codificador entrópico
TComLoopFilter	Deblocking filter
TComSampleAdaptiveOffset	Sample Adaptive Offset filter

Tabla 8 Clases del HM 8.0

Fichero	Localización	Descripción
TAppDecCfg	TAppDecoder	Configuración para la decodificación
TAppDecTop	TAppDecoder	Decodificador
TAppEncCfg	TAppEncoder	Configuración para la codificación
TAppEncTop	TAppEncoder	Codificador
TComTrQuant	TLibCommon	Transformación and cuantificación
TComPrediction	TLibCommon	Predicción
TComInterpolationFilter	TLibCommon	Filtro de interpolación
TComMv	TLibCommon	Vectores de movimiento
TComRdCost	TLibCommon	Coste de la tasa de distorsión (RD – Rate Distortion)
TComRdCostWeightPrediction	TLibCommon	Coste de la tasa de distorsión, con predicción con peso
TComWeightPrediction	TLibCommon	Predicción con pesos
WeightPredAnalysis	TLibEncoder	Análisis para la predicción con pesos
TComCABACTables	TLibCommon	Tablas para CABAC
TDecBinCoderCABAC	TLibDecoder	Decodificador entrópico binario de CABAC
TDecEntropy	TLibDecoder	Decodificador entrópico
TEncBinCoderCABAC	TLibEncoder	Codificador entrópico binario de CABAC
TEncEntropy	TLibEncoder	Codificador entrópico
TComLoopFilter	TLibCommon	Deblocking filter
TComSampleAdaptiveOffset	TLibCommon	Sample Adaptive Offset filter
TEncSampleAdaptiveOffset	TLibEncoder	Estimación empleada en el filtro SAO

Tabla 9 Ficheros del HM 8.0

3.3. Parámetros para la codificación

Para ejecutar el proceso de codificación en el software de referencia se debe ejecutar la siguiente línea de comando:

TAppEncoder [-h] [-c config.cfg] [--parameter=value]

Donde las opciones son las siguientes:

Opción	Descripción
-h	Imprime los parámetros a usar (help)
-c	Indica el fichero de configuración a usar. Se pueden emplear varios ficheros repitiendo la opción –c
--parameter=value	Asigna valores al parámetro descrito

En la Tabla 10 se recogen los parámetros, que en base al análisis realizado del software, se han considerado más representativos en el proceso de codificación. Los parámetros se han agrupado según su área de influencia. Si el parámetro no tiene "shorthand", la celda correspondiente aparece en blanco.

Parámetro	Shorthand	Descripción
Parámetros relacionados con I/O		
InputFile	-i	Especifica el fichero de vídeo de entrada
BitstreamFile	-b	Especifica el bitstream de salida
ReconFile		Especifica el fichero de las muestras reconstruidas
SourceWidth	-wdt	Especifica el ancho del vídeo de entrada en muestras de luma
SourceHeight	-hgt	Especifica el alto del vídeo de entrada en muestras de luma
InputBitDepth		Especifica la profundidad de bit empleada (por defecto es 8 bits por muestra)
FrameRate	-fr	Especifica la tasa de tramas del vídeo de entrada
FramesToBeEncoded	-f	Especifica el número de tramas a ser codificadas

Parámetro	Shorthand	Descripción
Parámetros relacionados con la estructura de los bloques		
MaxCUWidth		Define el ancho máximo para las CUs
MaxCUHeight		Define la altura máxima para las CUs
MaxCUSize		Define el tamaño máximo para las CUs
MaxPartitionDepth	-h	Define la profundidad del árbol de CUs
QuadtreeTULog2MaxSize		Define el tamaño máximo de las Tus en logaritmo base 2 (ejemplo: 6=log2(64))
QuadtreeTULog2MinSize		Define el tamaño mínimo de las Tus en logaritmo base 2
QuadtreeTUMaxDepthIntra		Define la profundidad del árbol de Tus para las CUs intra
QuadtreeTUMaxDepthInter		Define la profundidad del árbol de Tus para las CUs inter
Parámetro	Shorthand	Descripción
Parámetros relacionados con la estructura de codificación		
IntraPeriod	-ip	Especifica el período de las tramas intra (-1 indica que sólo la primera es codificada como intra)
DecodingRefreshType	-dr	Especifica el tipo de actualización de decodificación a aplicar al periodo de las imágenes I (intra): 0 Aplica una imagen I 1, 2 Aplica random Access
GOPSize	-g	Especifica el tamaño del GOP
Parámetro	Shorthand	Descripción
Parámetros relacionados con la estimación del movimiento		
FastSearch		Habilita o deshabilita el uso del “fast motion search”
SearchRange	-sr	Especifica el rango usado en la estimación de movimiento
BipredSearchRange		Especifica el rango usado para la bi-predicción en la estimación de movimiento
HadamardME		Habilita o deshabilita el uso de la transformada Hadamard en la estimación de las muestras sub-píxel
AMP		Habilita o deshabilita el uso de las particiones asimétricas

Parámetro	Shorthand	Descripción
Parámetros relacionados con la cuantificación		
QP	-q	Especifica el valor base de los parámetros de cuantificación
RDOQ		Habilita o deshabilita la optimización de los QPs basado en el ratio de distorsión
Parámetro	Shorthand	Descripción
Parámetros relacionados con los filtros		
LoopFilter Disable		Habilita o deshabilita el deblocking filter
SAO		Habilita o deshabilita el filtro SAO

Tabla 10 Parámetros principales en la codificación

Capítulo 4: Análisis funcional

En capítulos anteriores se han explicado las razones de la creación del estándar HEVC, se ha profundizado en los puntos que se han considerado de mayor interés y se ha presentado el código de referencia que emplearemos para el análisis. El propósito principal del HM, HEVC Test Model, es el de ofrecer una herramienta útil con la que explorar este nuevo estándar. El uso de este modelo nos ayudará a caracterizar el estándar. Se recomienda emplear las condiciones de test común las cuales incluyen cuatro modos de codificación predefinidos: all-intra (AI), random access (RA), low-delay P (LP) y low-delay B (LB).

4.1. Introducción

Algunos de los estudios previos han basado el análisis de los resultados obtenidos en HEVC basándose en medidas subjetivas de calidad tales como *Mean Option Score* (MOS). Sin embargo, existen medidas reproducibles para obtener la calidad objetivamente, por ejemplo *Peak Signal-to-Noise Ratio* (PSNR). Existen otras medidas para obtener un resultado objetivo pero ésta es la más simple y la más popular.

Con este trabajo se quiere proporcionar datos de calidad de manera objetiva. Por esta razón, y dada la sencillez y popularidad del método, se ha decidido emplear PSNR en este análisis. Las condiciones que se emplearán serán las recomendadas: all-intra (AI), random access (RA), low-delay P (LP) y low-delay B (LB). Además de ofrecer este dato, se caracterizarán las etapas de codificación teniendo en cuenta cuánto tiempo se invierte en cada una de ellas para cada una de las condiciones de test. Para ello, se empleará la herramienta Intel VTune Amplifier XE 2013. Todo ello se realizará empleando un procesador Intel Core i5 2450M.

El modelo de referencia a emplear será el HM 8.0, como ya se ha comentado con anterioridad. Las diferencias que se pudieran obtener en los resultados si usáramos el HM 11.0 (última versión) son en este caso marginales.

4.2. Condiciones de Test

Este experimento ha sido analizado bajo las configuraciones AI, RA, LP y LB usando la estructura adoptada principalmente en [11].

Para la condición AI, todas imágenes son codificadas como Intra (I) en el mismo orden en el que se muestran y sin referencias temporales ni offsets para los QPs.

Para la condición RA, imágenes tipo I son introducidas dependiendo del período asignado y el resto de imágenes son codificadas como tipo B (Inter bidireccionales). Esta configuración explota 4 capas (L1, L2, L3 y L4) usando una estructura jerárquica en la cual el GOP (Group Of Pictures) es 8. En la Figura 11 se pueden observar las primeras nueve imágenes empleadas en la secuencia, una imagen I y 8 imágenes B. El orden de codificación de las imágenes en el GOP es B_8, B_4, B_2, B_1, B_3, B_6, B_5, B_7. Las imágenes tipo I son codificadas con el QP original, para las tipo B, el QP tiene un offset igual a la capa/nivel en el que se encuentren. En la Figura 11 también se puede observar la lista de dependencia empleada en la predicción. Por ejemplo B_1 usa I, B_8, B_4 y B_2 como referencias.

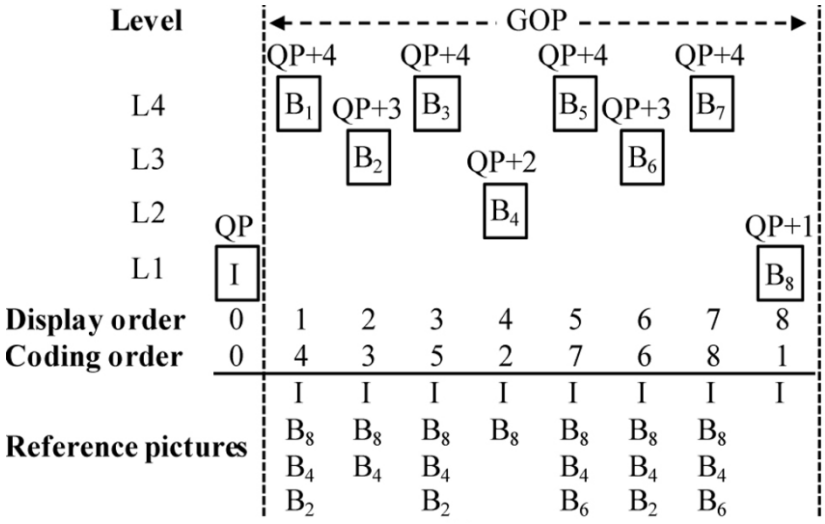

Figura 11 Estructura de codificación jerárquica del modo de configuración RA

Para la condición LB, sólo la primera imagen es tipo I y el resto de imágenes son codificadas como tipo B (Inter bidireccionales). Esta configuración explota 3 capas (L1, L2 y L3) usando una estructura jerárquica en la cual el GOP (Group Of Pictures) es 4. En la Figura 12 se pueden observar las primeras cinco imágenes empleadas en la secuencia, una imagen I y 4 imágenes B. El orden de codificación de las imágenes en el GOP es el mismo en el que se muestran, esto es B_1, B_2, B_3, B_4. Las imágenes tipo I son codificadas con el QP original (32), para las tipo B, el QP tiene un offset igual a la capa/nivel en el que se encuentren, al igual que en RA. En la Figura 12 también se puede observar la lista de dependencia empleada en la predicción. Por ejemplo B_1 usa I como referencia.

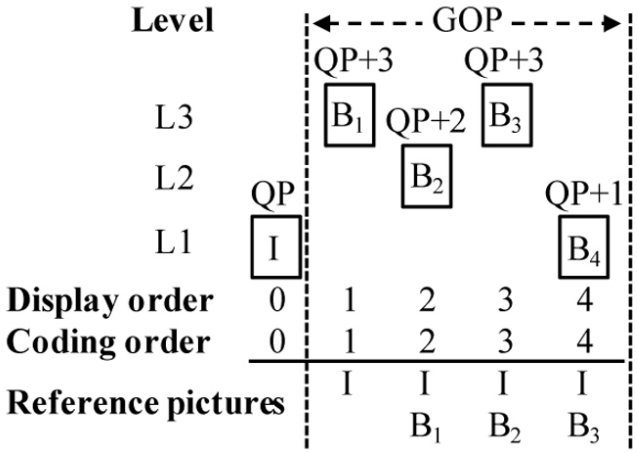

Figura 12 Estructura de codificación jerárquica del modo de configuración LB

La condición LP es similar a la LB pero cambiando las tramas B por P.

En el análisis se jugará con distintos QPs. Los valores elegidos han sido: 22, 27, 32 y 37.

4.3. Secuencias de Test

La Tabla 11 muestra las secuencias de 8 bits que se van a emplear en este análisis. Estas secuencias serán empleadas tanto para el análisis de PSNR como para el análisis de complejidad (en cuanto a tiempo empleado en el proceso) de los bloques.

Formato	Secuencia	Número de tramas	Frecuencia	Condiciones			
				AI	RA	LB	LP
832x480 (WVGA)	RaceHorses	300	30 fps	x	x	x	x
416x240 (WQVGA)	BlowingBubbles	500	50 fps	x	x	x	x
352x288 (CIF)	Mobile	300	30 fps	x	x	x	x
176x144 (QCIF)	Claire	494	30 fps	x	x	x	x

Tabla 11 Lista de secuencias de test

Como se puede observar en la Tabla 11, se van a emplear cuatro diferentes secuencias en este análisis. Las secuencias de test no han sido elegidas de manera aleatoria, se han seleccionado de forma que tengamos un amplio abanico de características. La secuencia de vídeo Claire consiste en una mujer dando una especie de telediario donde el fondo aparece en un único tono y de manera estática. La secuencia Mobile en cambio, es una secuencia con bastante movimiento. Se trata de un tren de juguete que sigue un recorrido a la vez que también encontramos movimiento en el fondo de la imagen. BlowingBubbles es un vídeo sobre dos chicas que están haciendo pompas de jabón. El fondo se puede considerar estático y las pompas de jabón van creciendo y moviéndose de manera aleatoria. Por último, la secuencia RaceHorses muestra una carrera de caballos. Al igual que en la secuencia Mobile, tanto el primer plano como el fondo están en movimiento. Como también se puede apreciar en la tabla anterior, el formato de cada una de las secuencias es diferente, abarcando desde QCIF a WVGA. El número de tramas y la frecuencia de las mismas también varían entre las distintas secuencias.

4.4. Plataforma de test

A continuación se indican las características de la plataforma empleada para este análisis. Durante el análisis, únicamente el software encargado de la codificación ha estado corriendo, de este modo se pretende evitar que los resultados sean afectados por otros procesos.

Procesador	Intel Core i5-2450M (2x2,5GHz)
Memoria	6,00 GB
Sistema operativo	64-bit Microsoft Windows 7 Home Premium SP1
Compilador	Microsoft Visual C++ 2008
Análisis de complejidad	Intel VTune Amplifier XE 2013

Para el análisis de complejidad contaremos con que el software VTune Amplifier. Esta herramienta de Intel proporciona información sobre el rendimiento del código. Para los desarrolladores software es una aplicación útil que les asiste en el análisis de rendimiento del código. De especial interés es la capacidad de proporcionar un análisis de los “hot spots” del proceso. Este software puede proporcionar más medidas, por ejemplo estimación de números de ciclos para cada una de las funciones del proceso. Para el estudio de complejidad desarrollado en este Trabajo Fin de Máster sólo se hará uso del análisis de “hot spots” identificando de esta manera qué etapas consumen más tiempo en el proceso de codificación.

Capítulo 5: Resultados y conclusiones

¿Cumplirá HEVC la meta de mejorar sustancialmente la tasa de bit manteniendo la misma calidad de imagen? ¿Vale la pena migrar a HEVC? Existen numerosas publicaciones que dan una respuesta positiva a estas preguntas, muchas de ellas basadas en MOS y otras en medidas objetivas como PSNR, tal y como se ha realizado en este trabajo. Es razonable mantener una aptitud escéptica a los resultados que se van a presentar en este trabajo ya que se ha usado un modelo de referencia y no un producto que esté en el mercado, pero los resultados aquí mostrados deben tomarse como indicadores consistentes de lo que podemos esperar en el mercado.

5.1. Introducción

Haciendo memoria de los objetivos principales de este Trabajo Fin de Máster, se han cubierto de manera satisfactoria los puntos relacionados con el estudio del estándar, comparativa con H.264 en cuanto a arquitectura planteada, estudio y familiarización con el modelo software de referencia, elección del setup para el análisis del estándar, tal y como se puede observar en capítulos anteriores.

En este capítulo, se presentarán los resultados que se han obtenido a base a las pruebas presentadas en el capítulo 4. Los resultados serán recogidos en tablas de manera que la presentación de los mismos sea clara y compacta.

Además, en este capítulo se presentarán las conclusiones obtenidas en base a los resultados y al mercado existente. También, se ha querido añadir a modo de "bonus" un pequeño resumen de las empresas y productos que están adaptando HEVC como estándar de codificación de vídeo.

5.2. Resultados

En la Tablas 12, 13, 14 y 15, se pueden observar los valores de bit rate y PSNR obtenidos en las cuatro condiciones de test especificadas anteriormente para los diferentes QPs en las distintas secuencias de vídeos empleadas. Los datos de bit rate son en kbits por segundo, mientras que las unidades de PSNR son decibelios (dB). En las tablas sólo se han reflejado los valores de PSNR de luminancia, ya que los valores de crominancia siguen la misma tendencia y no aportan información adicional a los resultados.

Secuencia	All Intra (AI)							
	22		27		32		37	
	BR (kbits/s)	PSNR (dB)	BR (kbits/s)	PSNR (dB)	BR (kbits/s)	PSNR (dB)	BR (kbits/s)	PSNR (dB)
Claire	244,326	45,864	155,378	42,284	97,062	38,563	59,583	35,162
Mobile	9480,275	41,465	6646,204	36,903	4429,964	32,423	2760,395	28,204
BlowingBubbles	11448,852	41,016	7070,036	36,712	4022,164	32,741	2070,618	29,202
RaceHorses	14982,426	42,278	8992,443	38,905	5105,663	35,492	2572,263	32,111

Tabla 12 Bit rate y PSNR obtenidos en AI

Secuencia	Random Access (RA)							
	22		27		32		37	
	BR (kbits/s)	PSNR (dB)	BR (kbits/s)	PSNR (dB)	BR (kbits/s)	PSNR (dB)	BR (kbits/s)	PSNR (dB)
Claire	23,969	44,933	13,410	41,715	7,931	38,393	5,015	35,293
Mobile	1489,257	38,154	702,460	34,156	353,760	30,762	191,407	27,667
BlowingBubbles	1706,156	38,252	798,816	34,889	382,468	31,717	181,926	28,806
RaceHorses	4787,172	39,040	2024,750	35,8428	943,847	32,978	462,783	30,336

Tabla 13 Bit rate y PSNR obtenidos en RA

Secuencia	Low Delay B (LB)							
	22		27		32		37	
	BR (kbits/s)	PSNR (dB)	BR (kbits/s)	PSNR (dB)	BR (kbits/s)	PSNR (dB)	BR (kbits/s)	PSNR (dB)
Claire	22,391	43,956	11,200	40,423	6,077	37,157	3,612	33,991
Mobile	1761,144	38,442	744,438	34,008	316,910	30,279	140,806	26,799
BlowingBubbles	1928,777	38,077	821,916	34,315	352,058	30,837	149,270	27,772
RaceHorses	5694,716	39,877	2261,156	36,203	995,037	33,032	456,140	30,162

Tabla 14 Bit rate y PSNR obtenidos en LB

Secuencia	Low Delay P (LP)							
	22		27		32		37	
	BR (kbits/s)	PSNR (dB)	BR (kbits/s)	PSNR (dB)	BR (kbits/s)	PSNR (dB)	BR (kbits/s)	PSNR (dB)
Claire	24,150	43,899	11,529	40,408	6,124	37,173	3,717	34,166
Mobile	2102,471	38,340	852,971	33,728	345,060	30,103	146,282	26,795
BlowingBubbles	2138,416	37,965	859,297	34,226	360,211	30,826	151,527	27,788
RaceHorses	6057,248	39,832	2359,126	36,069	1018,564	32,923	462,492	30,097

Tabla 15 Bit rate y PSNR obtenidos en LP

En las Figuras 13, 14, 15 y 16, se pueden observar gráficamente los resultados obtenidos para cada secuencia por separado.

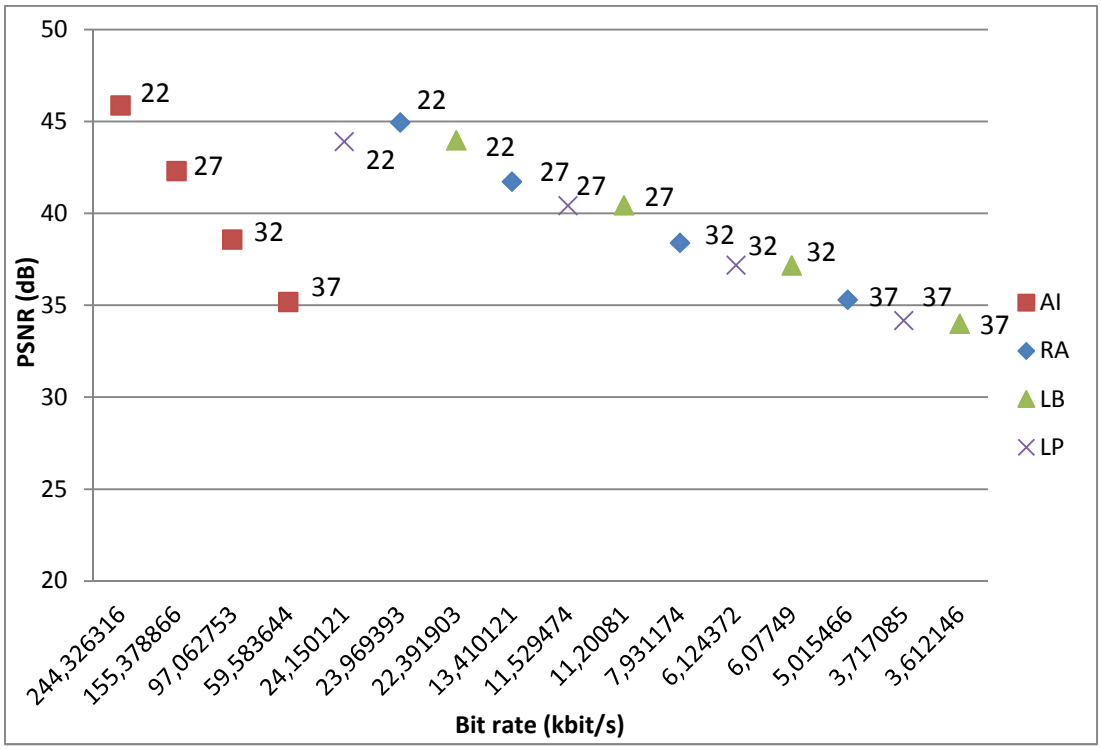

Figura 13 Resultados para la secuencia Claire

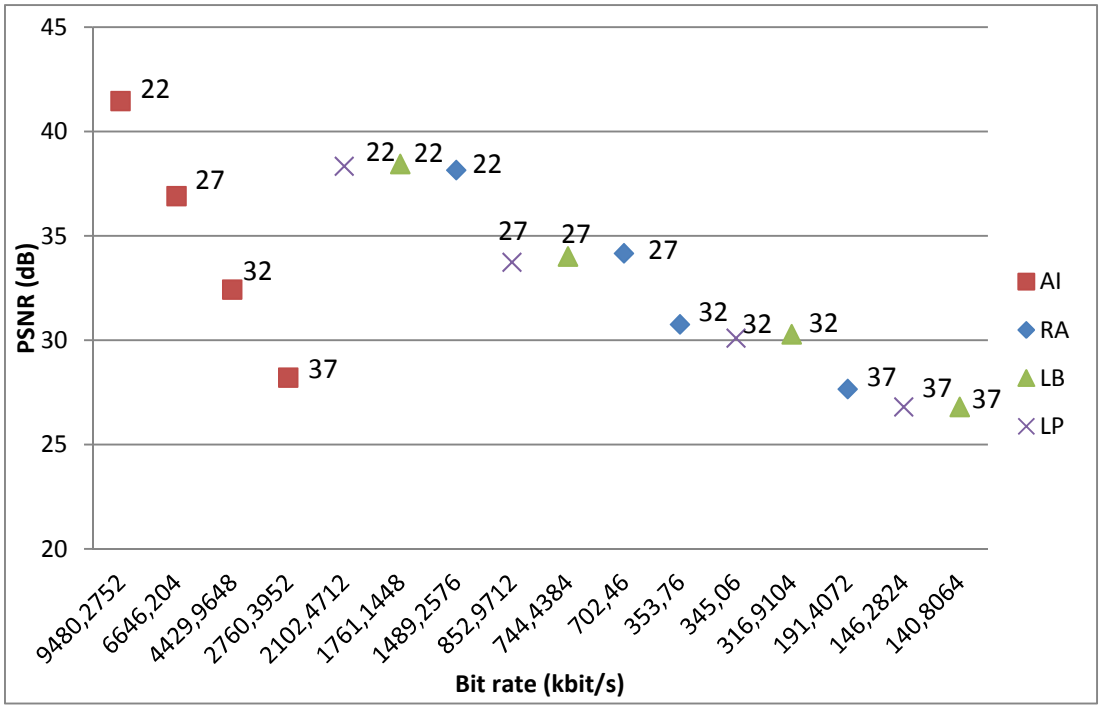

Figura 14 Resultados para la secuencia Mobile

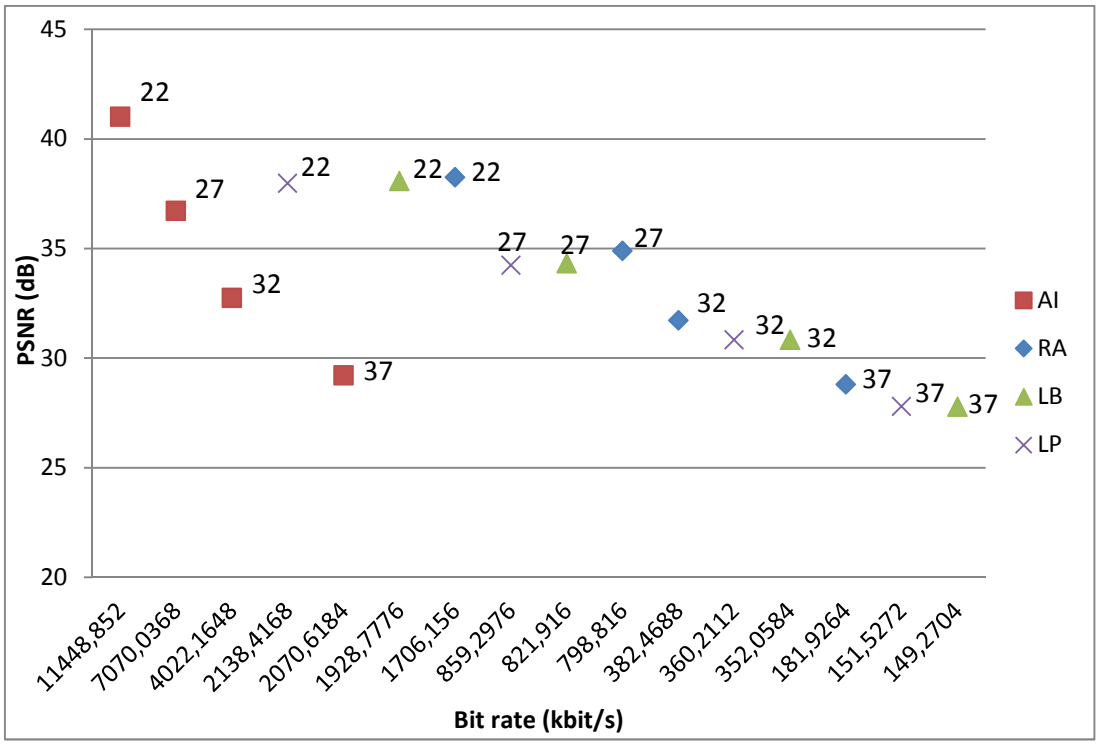

Figura 15 Resultados para la secuencia BlowingBubbles

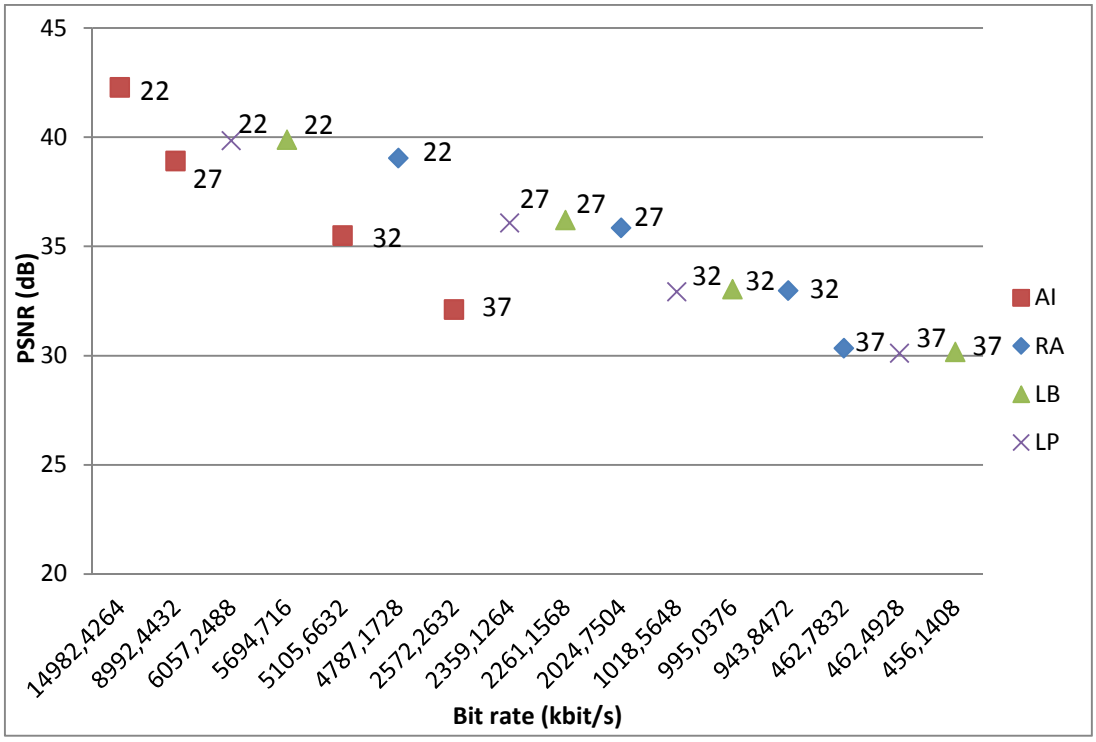

Figura 16 Resultados para la secuencia RaceHorses

Si analizamos los datos, comprobamos como para todas las secuencias y condiciones de test, obtenemos una reducción de la tasa de bit, a costa de un decremento del PSNR, a medida que se incrementa el QP. Codificando todas las tramas con predicción espacial (AI) obtenemos mejor calidad de imagen pero el bit rate necesario es más de 9 veces mayor si lo comparamos con el resultado obtenido en el resto de configuraciones con un sacrificio del PSNR que se podría considerar bajo, entre 1 y 3 dB. Por otra parte, Random Access obtiene mejores resultados que Low Delay B y este, a su vez, mejores resultados que Low Delay P. El motivo de obtener mejores resultados en Random Access se debe a que con esta configuración se introduce una imagen tipo I cada 32 imágenes, a este tipo de tramas se les aplica un QP menor, por lo cual, aumentamos la calidad final. Y finalmente, Low Delay B obtiene mejores resultados que Low Delay P ya que las imágenes tipo B obtienen una mejor tasa de compresión que las de tipo P, por tanto, a igual bit rate la calidad es mayor en Low Delay B, o lo que es lo mismo, a igual calidad, menor bit rate es necesario.

En la Tablas 16, 17, 18 y 19, se muestran los datos obtenidos en el análisis de carga computacional realizado con la herramienta Intel VTune Amplifier. Los datos han sido recogidos en base a las funciones de codificación que se han considerado más relevantes en este estudio, más concretamente aquellas relacionadas con la estimación de movimiento, predicción intra/inter, transformación-cuantificación, codificación entrópica y loop filters. Los valores presentados en las tablas representan el porcentaje empleado por cada función en el proceso de codificación. En las tablas presentadas abajo sólo se mostrarán valores relativos al total empleado en cada una de las simulaciones.

Secuencia		All Intra (AI)									
		Inter Filter (%)	RD Cost (%)	IQ/TQ (%)	memcpy (%)	Sbac (%)	Cabac cnt (%)	Pred (%)	Entropy (%)	SAO+DF (%)	Otros (%)
Claire	22	0	8,156	34,381	8,156	8,736	2,523	7,155	1,138	0,675	29,081
	27	0	7,803	32,063	8,731	7,575	2,062	6,773	1,175	0,728	33,090
	32	0	8,593	29,767	9,099	6,453	1,730	7,188	1,228	0,779	35,163
	37	0	9,061	27,641	9,778	5,849	1,266	7,609	1,117	0,882	36,797
Mobile	22	0	4,830	42,624	5,626	13,715	4,574	4,210	0,515	0,421	23,485
	27	0	5,290	41,276	5,983	12,789	4,298	4,537	0,518	0,468	24,841
	32	0	5,903	39,959	6,533	11,384	3,668	5,002	0,556	0,505	26,490
	37	0	5,616	37,726	7,201	9,732	2,990	5,616	0,599	0,578	29,942
BlowingBubbles	22	0	4,911	42,083	5,746	13,059	4,302	4,350	0,505	0,453	24,591
	27	0	5,704	39,933	6,475	11,248	3,603	4,887	0,554	0,520	27,076
	32	0	6,500	36,418	7,296	8,975	2,674	5,558	0,590	0,609	31,380
	37	0	7,625	33,072	8,236	6,851	1,870	6,442	0,644	0,705	34,555
RaceHorses	22	0	6,372	39,533	6,841	8,772	3,308	5,308	0,549	0,536	28,780
	27	0	7,307	36,194	7,663	8,902	2,666	6,037	0,601	0,618	30,013
	32	0	8,223	33,223	8,492	7,318	2,048	6,718	0,645	0,696	32,637
	37	0	9,201	30,149	9,228	5,862	1,464	7,396	0,677	0,776	35,246

Tabla 16 Distribución de complejidad para All Intra

Secuencia		Random Access (RA)									
		Inter Filter (%)	RD Cost (%)	IQ/TQ (%)	memcpy (%)	Sbac (%)	Cabac cnt (%)	Pred (%)	Entropy (%)	SAO+DF (%)	Otros (%)
Claire	22	31,906	25,373	14,351	3,965	2,831	0,784	0,755	0,459	0,159	19,416
	27	33,961	27,037	11,797	4,069	2,037	0,505	0,737	0,441	0,173	19,243
	32	34,844	28,052	10,049	4,145	1,646	0,380	0,718	0,431	0,185	19,549
	37	35,207	28,766	9,079	4,127	1,466	0,258	0,677	0,405	0,180	19,835
Mobile	22	23,439	17,604	26,068	3,574	7,567	2,747	0,987	0,547	0,096	17,371
	27	28,669	21,065	20,977	3,563	4,991	1,679	0,835	0,499	0,113	17,609
	32	33,048	23,935	16,138	3,493	3,160	0,995	0,739	0,458	0,143	17,891
	37	35,706	25,661	12,927	3,541	2,184	0,611	0,705	0,423	0,159	18,083
BlowingBubbles	22	25,300	19,955	23,780	3,570	6,039	2,084	0,913	0,521	0,102	17,736
	27	29,746	23,177	18,901	3,630	3,780	1,216	0,850	0,485	0,125	18,090
	32	32,960	25,584	14,654	3,713	2,482	0,696	0,769	0,457	0,143	18,542
	37	34,998	27,338	11,711	3,761	1,770	0,422	0,713	0,416	0,168	18,703
RaceHorses	22	21,339	27,132	21,726	3,187	6,238	2,183	0,926	0,520	0,090	16,659
	27	25,782	30,911	16,902	3,237	3,744	1,211	0,812	0,486	0,100	16,815
	32	29,215	32,587	13,736	3,229	2,433	0,713	0,723	0,458	0,096	16,811
	37	31,911	33,254	11,400	3,220	1,748	0,450	0,665	0,434	0,115	16,802

Tabla 17 Distribución de complejidad para Random Access

Secuencia		Low Delay B (LB)									
		Inter Filter (%)	RD Cost (%)	IQ/TQ (%)	memcpy (%)	Sbac (%)	Cabac cnt (%)	Pred (%)	Entropy (%)	SAO+DF (%)	Otros (%)
Claire	22	32,583	28,007	14,089	3,258	2,499	0,748	0,646	0,435	0,117	17,618
	27	35,002	29,848	10,968	3,273	1,686	0,440	0,567	0,379	0,094	17,743
	32	36,685	31,478	8,682	3,224	1,312	0,284	0,508	0,313	0,117	17,397
	37	36,783	32,592	7,551	3,185	1,137	0,225	0,438	0,325	0,155	17,610
Mobile	22	24,420	19,803	24,536	3,125	7,507	2,740	0,894	0,490	0,075	16,410
	27	29,628	23,459	19,914	2,996	4,801	1,658	0,745	0,459	0,078	16,261
	32	34,081	26,808	15,159	2,903	2,852	0,912	0,619	0,398	0,067	16,201
	37	36,761	29,121	11,851	2,845	1,728	0,489	0,528	0,360	0,069	16,249
BlowingBubbles	22	26,261	22,578	22,359	3,117	5,690	1,996	0,832	0,467	0,074	16,626
	27	30,853	26,591	17,228	3,089	3,299	1,067	0,734	0,429	0,077	16,632
	32	33,807	28,873	13,674	3,049	2,091	0,605	0,627	0,392	0,080	16,802
	37	35,016	30,185	10,474	2,979	1,386	0,338	0,554	0,364	0,087	18,619
RaceHorses	22	22,812	32,538	17,943	2,766	5,057	1,754	0,800	0,429	0,071	15,830
	27	26,065	35,878	14,339	2,674	3,183	1,028	0,700	0,405	0,077	15,651
	32	29,203	37,621	11,576	2,579	1,998	0,587	0,602	0,375	0,076	15,384
	37	31,869	38,322	9,534	2,558	1,353	0,348	0,536	0,352	0,066	15,063

Tabla 18 Distribución de complejidad para Low Delay B

Secuencia		Low Delay P (LP)									
		Inter Filter (%)	RD Cost (%)	IQ/TQ (%)	memcpy (%)	Sbac (%)	Cabac cnt (%)	Pred (%)	Entropy (%)	SAO+DF (%)	Otros (%)
Claire	22	25,669	21,678	21,086	4,859	3,841	1,155	0,885	0,614	0,169	20,044
	27	29,045	24,120	16,683	4,998	2,568	0,680	0,739	0,577	0,141	20,449
	32	31,789	25,733	13,388	4,936	2,040	0,438	0,650	0,517	0,178	20,333
	37	32,625	26,630	11,675	4,938	1,687	0,339	0,603	0,527	0,235	20,740
Mobile	22	15,258	13,584	33,076	4,065	10,819	3,958	1,054	0,569	0,095	17,523
	27	20,295	17,026	28,754	4,187	7,561	2,653	0,933	0,623	0,120	17,848
	32	26,004	20,785	22,907	4,287	4,609	1,523	0,776	0,572	0,135	18,402
	37	30,077	23,864	18,116	4,302	2,814	0,803	0,673	0,539	0,114	18,698
BlowingBubbles	22	17,693	17,625	29,888	4,218	7,831	2,816	0,998	0,632	0,113	18,186
	27	22,668	20,978	24,924	4,358	5,085	1,617	0,905	0,630	0,127	18,708
	32	26,652	23,995	20,102	4,509	3,213	0,922	0,813	0,585	0,129	19,081
	37	29,694	26,230	16,101	4,604	2,153	0,535	0,707	0,566	0,130	19,282
RaceHorses	22	14,894	31,308	22,776	3,489	6,720	2,353	0,905	0,537	0,090	16,928
	27	17,985	35,119	19,005	3,496	4,416	1,447	0,806	0,534	0,100	17,092
	32	21,230	37,459	15,854	3,511	2,842	0,852	0,710	0,506	0,111	16,924
	37	24,216	38,315	13,359	3,561	1,960	0,509	0,640	0,490	0,096	16,854

Tabla 19 Distribución de complejidad para Low Delay P

Agrupando las funciones según a la fase de codificación a la que pertenecen (memcpy se encarga de accesos a memoria, por tanto, a ninguna fase en particular), y haciendo la media obtenemos los resultados que se recogen en la Tabla 20.

Fase de codificación	AI	RA	LB	LP
Interpolación Inter	0%	55%	59%	57%
Interpolación Intra	15%	1%	1%	1%
T/Q/IT/IQ	37%	15%	14%	20%
EC	11%	4%	3%	5%
Misceláneos	37%	25%	23%	17%

Tabla 20 Media de las fases de codificación con mayor complejidad

La codificación empleando AI es la de menor complejidad de todas ya que no emplea interpolación temporal (interpolación inter). Debido a la no existencia de este tipo de interpolación, la carga computacional requerida en la etapa de transformación/cuantificación es mayor en el caso de AI que en el resto de las condiciones de test. Siendo para RA, LB y LP un 59.45%, 62.16% y 45.94% menor que en AI, respectivamente. De la misma manera, en el caso de AI, el tiempo empleado en la codificación entrópica es mayor que en el resto de condiciones de test. La inclusión de la predicción temporal incrementa la complejidad en los casos RA, LB y LP, aumentando el tiempo de ejecución total de la codificación en 2.42 el tiempo empleado en AI (un incremento de aproximadamente un 141%). Las etapas de transformación/cuantificación y codificación entrópica reducen su complejidad como ya se ha comentado previamente. El valor de QP, al igual que en los resultados anteriores de calidad, tiene un impacto en el tiempo empleado para la codificación. Dependiendo de la condición de test y la secuencia empleada, incrementando el QP de 22 a 27 la reducción oscila entre el 7% y 16%, yendo de 27 a 32 la reducción se encuentra entre el 10% y el 21% y, por último, si se incremente el QP de 32 a 37 la reducción está en el rango del 9% al 21%.

5.3. Conclusiones

El estándar HEVC ha sido desarrollado conjuntamente por la ISO/IEC y la ITU-T, en conjunto JCT-VC. HEVC es el sucesor del H.264/AVC, estándar que actualmente se emplea en productos como Adobe, Apple, France Telecom, Intel, Microsoft, Motorola, Nokia, Samsung, Sony, Toshiba y otras empresas, para transmitir vídeo de alta definición así como en formatos de disco Blu-Ray, teléfonos móviles, herramientas de videoconferencia, medios de almacenamiento digital, etc. A día de hoy, H.264/AVC es la norma de compresión de vídeo más utilizada en el mundo.

En este Proyecto Fin de Máster se ha presentado el estudio realizado sobre el nuevo estándar de codificación de vídeo, el estándar HEVC, así como resultados sobre su rendimiento y coste computacional. Las simulaciones se han realizado para las condiciones de test AI, RA, LB y LP empleando QP 22, 27, 32 y 37. Además, se han empleado secuencias de diferente naturaleza, con resoluciones que varían desde QCIF hasta WVGA y empleando diferente número de tramas y frecuencia.

HEVC ofrece una solución flexible, fiable y robusta, representando a la vez una serie de avances en la codificación de vídeo. Es un codificador de vídeo basado en la misma filosofía que estándares previos, es decir, es un codificador de vídeo híbrido, pero aporta algunas importantes ventajas con respecto a estándares anteriores, como ya se describió en el capítulo 2 de este Trabajo Fin de Máster. La complejidad añadida debida a las nuevas modificaciones es afrontable aunque, en base a los resultados obtenidos, se requerirá trabajar duramente en la implementación de la estimación de movimiento, en particular, en dispositivos móviles donde el consumo de potencia y el tamaño del chip son fundamentales. Aunque quede fuera de los objetivos de este Trabajo Fin de Máster hacer una comparación de los resultados entre HEVC y H.264, me gustaría apuntar que estas nuevas características proporcionan aproximadamente un ahorro del 50% de la tasa de bit para obtener la misma calidad de imagen que su predecesor [11].

5.4.Mercado

El vídeo digital es cada día más importante, un estudio reciente realizado por Rovi Corporation muestra que aproximadamente el 63% de los usuarios de tabletas y el 71% de los usuarios de teléfonos móviles en EE.UU. visualizaron vídeos en esos dispositivos de 2 a 3 veces por semana o más. Al mismo tiempo, los usuarios de dispositivos móviles destacaron estar preocupados por el coste relacionado con la transmisión de vídeo. Aproximadamente la mitad de los encuestados afirmó que limitan la cantidad de streaming de vídeo en sus teléfonos móviles debido a la preocupación por el coste derivado del mismo. Si a esto sumamos, la lentitud que se observa en los tiempos de carga de vídeo, podemos extraer que desde el punto de vista del usuario, HEVC sería acogido gratamente.

Los usuarios de servicios cloud también se verían beneficiados de la implantación de HEVC en el mercado ya que muchas empresas que ofrecen este servicio están cobrando una cuota por todo lo que no es el paquete básico, por lo cual los consumidores tienen que restringir considerablemente lo que ponen en la nube. Con HEVC, los consumidores podrán almacenar muchos más vídeos y fotos por el mismo precio.

Otro gran impulsor del HEVC, es el vídeo de súper alta resolución, vídeos 4K. El tamaño de estos vídeos es enorme pero con HEVC su auge puede ser posible. El formato 4K podría ser el gran aliado de HEVC, como lo fue HD de H.264/AVC.

Con lo expuesto anteriormente parece obvio que las empresas proveedoras de codecs de vídeo quieran adoptar este estándar lo antes posible para poder ser realmente competitivos en este mercado. Es por ello por lo que HEVC ha sido este año la estrella de numerosas ferias en el mundo. En la Tabla 21 se presentan algunas de las implementaciones y productos HEVC.

Empresa	Implementación/Producto
Qualcomm	Decodificador HEVC corriendo en una tableta Android.
Ericsson	Ericsson SVP 5500. Codificador HEVC diseñado para soportar tiempo real en dispositivos móviles.
Allegro DVT	IP para la decodificación de vídeo. Esta IP puede ser usada en FPGA o SoC soportando resoluciones hasta 4K.
MIT	Presentó el primer decodificar HEVC implementado en ASIC capaz de decodificar un vídeo 3840x2160 a 30fps en tiempo real.
ATEME	Primer reproductor open source software para HEVC.
Broadcom	BCM7445. Decodificador Ultra HD. Basado en arquitectura ARM. Producción prevista para mediados de 2014.
Elemental Technologies	HEVC/H.265 encoder. Soporte mediante actualizaciones software.
Envivio	Envivio Muse. Codificador software HEVC. HEVC Early Access Program para aplicaciones bajo demanda.
Vanguard Video	Solución puramente software. Codificador HEVC en tiempo real.
DivX	DivX HEVC. Soporta Main Profile con algunas restricciones, soporta 4K, 1080p y 720p. Decoder plugin para DivX Player. DivX Web Player soportará HEVC.
Imagination Technologies	PowerVR Series5 HEVC decoder.
Fraunhofer HHI	HEVC 4K real-time decodificador software. HEVC 1080p HD real-time software codificador. HEVC 1080p HD real-time decodificador hardware.
Microsoft	Presenta las especificaciones de DirectX Video Acceleration (DXVA) para HEVC que soporten Main Profile, Main 10 Profile y Main Still Picture profile.
ViXS Systems	XCode 6400 SoC. Soporta resolución 4K a 60fps y Main 10 profile.
Samsung Electronics	Afirma que la serie de TVs de plasma F8500 y los Samsung Galaxy S4 soportarán HEVC.
LG	Afirma que la serie LA9700 (4K Ultra HD TV) soportará decodificación HEVC.

Tabla 21 Implementaciones y productos HEVC

Bibliografía

[1] Line transmission on non telephone signals. Video Codec for audiovisual services at 64 Kbits/s. ITU-T Recommendation H.261, Marzo de 1993.

[2] ISO/IEC JTC1 CD 11172 “Coding of moving pictures and associated audio for digital storage media up to 1.5 Mbp/s”. ISO 1992.

[3] ISO/IEC 13818-2: “Information technology - Generic coding of moving pictures and associated audio information”. 2000.

[4] ISO/IEC 14496-2 “Information technology - Coding of Audio Visual Objects. Part 2: Visual”. 2001.

[5] Recommendation ITU-T H.263 “Video Coding for Low Bit-Rate Communication”, version 1, 1995.

[6] Draft ITU-T recommendation and final draft international standard of joint video specification. Joint Video Team (JVT) de ISO/IEC MPEG e ITU-T VCEG, 2003.

[7] Recommendation ITU-T H.265 “High efficiency video coding”, 2013.

[8] High Efficiency Video Coding: The Net Frontier in Video Compression. Jens-Rainer Ohm and Gary J. Sullivan, December 2012.

[9] Fraunhofer Heinrich-Hertz-Institut, website: http://hevc.hhi.fraunhofer.de/

[10] HM HEVC Reference Software v8.0, en línea, Disponible en: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-8.0/

[11] Comparative Rate-Distortion-Complexity Analysis of HEVC and AVC Video Codecs. Jarno Vanne, Marko Viitanen, Timo D. Hämäläinen, Antti Hallapuro. December 2013.