



Secretaría de Estado de Telecomunicaciones
y para la Sociedad de la Información



Foro de la TV de Alta Definición

FORMATOS DE PRODUCCIÓN, INTERCAMBIO Y DIFUSIÓN DE CONTENIDOS DE TV EN ALTA DEFINICIÓN

Versión 1.0

Elaborado por

**Grupo Técnico del Foro de la Televisión
de Alta Definición en España**

Coordinado por
Corporación RTVE

Abril de 2008

Índice

1	OBJETIVOS.....	3
2	ESTÁNDARES Y FORMATOS DE PRODUCCIÓN DE CONTENIDOS.....	4
2.1	NORMATIVA PARA LA PRODUCCIÓN Y EL INTERCAMBIO DE PROGRAMAS EN TVAD.....	5
2.2	NORMATIVA UIT-R EN TVAD	6
2.3	NORMATIVA SMPTE EN TVAD	7
2.4	RECOMENDACIONES EBU EN TVAD	10
2.5	PRODUCCIÓN DE AUDIO MULTICANAL	11
2.6	CONCLUSIONES	12
3	FORMATOS DE ALMACENAMIENTO E INTERCAMBIO DE PROGRAMAS.....	13
3.1	FORMATOS DE ALMACENAMIENTO E INTERCAMBIO DE VÍDEO	13
3.1.1	<i>MPEG-2 Intercuadro.....</i>	14
3.1.2	<i>AVC-Intra.....</i>	14
3.1.3	<i>JPEG2000.....</i>	15
3.1.4	<i>Otros formatos alternativos.....</i>	15
3.2	FORMATOS DE ALMACENAMIENTO E INTERCAMBIO DE AUDIO MULTICANAL.....	16
3.3	RECOMENDACIONES	17
4	FORMATOS DE CONTRIBUCIÓN DE CONTENIDOS.....	18
4.1	MPEG-2	18
4.2	MPEG-4 AVC/H.264	19
4.3	JPEG2000	19
4.4	RECOMENDACIONES	20
5	FORMATOS DE DIFUSIÓN DE CONTENIDOS.....	21
5.1	ESTÁNDARES DE CODIFICACIÓN DE VÍDEO	21
5.1.1	<i>MPEG-4 AVC/H.264</i>	21
5.1.2	<i>Otras tecnologías de codificación alternativas.....</i>	22
5.2	ESTÁNDARES DE CODIFICACIÓN DE AUDIO MULTICANAL.....	23
5.2.1	<i>AC3, E-AC3.....</i>	23
5.2.2	<i>HE-AAC v2.....</i>	24
5.3	RECOMENDACIONES	25

1 OBJETIVOS

El presente documento redactado para el Foro de la Televisión de Alta Definición, pretende reunir el estado del arte en TVAD¹ en sus aspectos técnicos relacionados con los formatos de producción, el archivo e intercambio de contenidos, así como la emisión de servicios de TV en Alta Definición.

Para ello se han sintetizado el conjunto de normativas definidas por los distintos organismos internacionales de estandarización que especifican dichos formatos. La normalización de formatos es un factor clave en el sector audiovisual, considerándose esta imprescindible para la correcta interoperabilidad entre los distintos agentes del sector, productores, operadores, radiodifusores y usuarios.

El grupo de trabajo se ha fijado como principal objetivo el realizar, de modo consensuado entre los participantes, la recomendación de los formatos de Alta Definición que se consideran óptimos en las etapas de producción, intercambio y difusión; de modo que se facilite un escenario armonizado para la difusión de estos servicios en España.

¹ Televisión de Alta Definición

2 ESTÁNDARES Y FORMATOS DE PRODUCCIÓN DE CONTENIDOS

2

El objetivo de este apartado es el de describir la actual normativa que define y regula los distintos formatos de producción de TV en Alta Definición, facilitando la interoperabilidad entre dispositivos y permitiendo un efectivo intercambio de contenidos audiovisuales entre los distintos actores que componen la cadena de producción.

Dichos formatos definen básicamente la resolución espacio-temporal de las secuencias de vídeo, y son definidos habitualmente por el número de líneas visibles que presentan, existiendo en la actualidad dos familias de formatos conocidas como **1080** y **720**.

Ambos formatos son una evolución de la Recomendación **UIT-R BT.601** [1], conocida durante muchos años como "Rec. 601", y que ha cumplido recientemente 25 años de vigencia. La BT.601 fue el primer estándar para vídeo digital en el entorno de producción, y ha sido adoptado por la totalidad de estándares de codificación de vídeo, como los definidos por la familia de estándares MPEG.

El formato 1080 nació con un doble objetivo, duplicar la resolución horizontal del actual formato de definición estándar de 720 píxeles/línea, y el de mejorar la relación de aspecto pasando de 4:3 a un formato panorámico de 16:9. Esto implica una resolución horizontal de 1920 píxeles/línea, que unido a la necesidad de definir una relación de píxel cuadrado, converge en el actual formato de 1080 líneas activas. Este formato es definido en Europa como **1080i25**, identificando su carácter entrelazado de 25 cuadros por segundo.

Por el contrario, el formato 720 perseguía el duplicar la resolución temporal de 25 cuadros/s a 50 cuadros/s pero en este caso con barrido progresivo, con el objetivo de adaptarse de modo óptimo a los contenidos con escenas de elevado movimiento, como ocurre con los contenidos deportivos. Al igual que el formato 1080, el formato 720 ofrece una mejora en su resolución espacial con respecto a la definición estándar, pero con una resolución horizontal de 1280 píxeles/línea ligeramente inferior a la del formato 1080; obtenida como consecuencia de aplicar un factor multiplicativo de 16/9 a la resolución horizontal del actual formato de definición estándar. Este formato es definido en Europa como **720p50**.

Las bondades del barrido progresivo, utilizado desde sus orígenes en los entornos cinematográficos, generaron una variante del formato 1080 con idéntica resolución espacial pero con 25 cuadros progresivos, definido como **1080p25**. Su resolución temporal convierte a este formato en el preferido para muchos creadores de géneros dramáticos, TV-Movies y algunos tipos específicos de documentales.

La razón por la que no se optó por definir en sus inicios un formato de 1080 líneas con barrido progresivo de 50 cuadros, fueron las dificultades tecnológicas para ser transmitido por los interfaces digitales serie basados en cable coaxial, al requerir de unas tasas binarias próximas a los 3Gbps.

En la actualidad esta barrera ha sido superada y en 2005 se estandarizó dicho formato denominado **1080p50**, considerándose como un primer paso para la **TVAD de 3^a generación**. Este formato es el preferido por la industria como formato de producción para la TVAD, ya que permite mantener toda la cadena de producción-difusión-visualización con el mismo barrido progresivo y duplicando la resolución temporal de los actuales formatos entrelazados. El barrido progresivo es el formato nativo de los dispositivos de captación y presentación digital, y eliminan de modo significativo las pérdidas de calidad debido a las conversiones de formatos entrelazado-progresivo.

A pesar de que a día de hoy la penetración comercial del formato 1080p50 es todavía limitada, ya se han comenzado a comercializar dispositivos de conmutación, distribución, captación y almacenamiento, y se estima que de modo progresivo se irán introduciendo el resto de elementos de la cadena como dispositivos de edición, postproducción, contribución y recepción. Probablemente todavía será necesario un plazo aproximado de 5 años para que este nuevo formato se consolide en el mercado.

2.1 Normativa para la producción y el intercambio de programas en TVAD

Desde 1990 se han aprobado una gran variedad de recomendaciones relativas a los distintos parámetros de producción de programas de Televisión en Alta Definición, por parte de los dos principales organismos de normalización de TV, la UIT (Unión Internacional de Telecomunicaciones) y el SMPTE (*Society of Motion Picture and Television Engineers*).

En sus inicios, estas normas definían parámetros de Alta Definición en un entorno de TV analógica con barrido entrelazado, debido a las limitaciones tecnológicas del momento, que unido al elevado coste requerido para su distribución, provocó el fracaso comercial generalizado de la TV de Alta Definición analógica.

El emergente éxito de la actual TV de Alta Definición digital, denominada **TVAD de 2^a generación**, se debe en gran medida a 2 factores:

1. El consenso a nivel mundial en la definición de un único formato de Alta Definición denominado **CIF²**, con una resolución vertical de la imagen de **1080** líneas visibles, rompiendo las diferencias históricas existentes en la TV de definición estándar (625 líneas y 525 líneas). Dicho formato se recoge en la Recomendación **UIT-R BT.709 Parte 2**[2], y ha sido ampliado y corregido en distintas ocasiones hasta su versión definitiva del 2002.
2. La evolución de la tecnología de compresión digital que facilita el almacenamiento, contribución y distribución de la TVAD con un coste muy similar a los requeridos para la TV de definición estándar, como es el caso del actual estándar **MPEG-4 AVC/H.264** [3].

En la actualidad, se han aprobado la práctica totalidad de estándares de Alta Definición para los entornos profesionales de producción, equivalentes a los definidos para la SDTV. En concreto se ha adoptado la misma topología de red punto a punto para la conexión de los distintos elementos de la cadena de producción, definiendo como interfaz de comunicación digital una evolución del interfaz SDI (*Serial Digital Interface*), denominado en este caso **HD-SDI** (*High Definition Serial Digital Interface*) y aprobados por la UIT como Recomendación BT.1120 [4], y por el SMPTE como Recomendación SMPTE 292M [5].

Ambos interfaces utilizan como medio físico el cable coaxial, con distancias máximas aproximadas de 300m para SD y 150m para HD, permiten el transporte de la señal de vídeo digital sin comprimir con regímenes binarios de 270Mbps y 1,485 Gbps respectivamente. La similitud entre los interfaces SDI y HD-SDI permite una fácil y rápida migración de los centros de producción de definición estándar a la TVAD.

El UIT aprobó una revisión del interfaz HD-SDI en su recomendación UIT-R BT.1120-5, que amplia el régimen binario de dicho interfaz hasta los 2,97Gbps, con el fin de poder integrar en un único enlace (modo Single-Link) los formatos CIF progresivos con frame

² Common Image Format

rates de 50Hz y 60Hz (1080p50 y 1080p60). Similares pasos ha dado el SMPTE con la estandarización de dicho interfaz como SMPTE 424M [6].

En el entorno multimedia también se ha definido el encapsulado de contenidos TVAD en formato de ficheros informáticos, basados en la sintaxis del nuevo estándar **MXF**³, estandarizado como SMPTE 377M [7]. En concreto, el SMPTE 384M [8] permite el intercambio electrónico de ficheros con streams de Alta Definición sin comprimir.

2.2 Normativa UIT-R en TVAD

Como se ha descrito anteriormente, los inicios de la TV en Alta Definición se deben en gran medida a la aprobación de la Recomendación **UIT-R BT.709** que abarca tanto los parámetros colorimétricos como los correspondientes a la digitalización de las componentes que definen el espacio de color del sistema de TV.

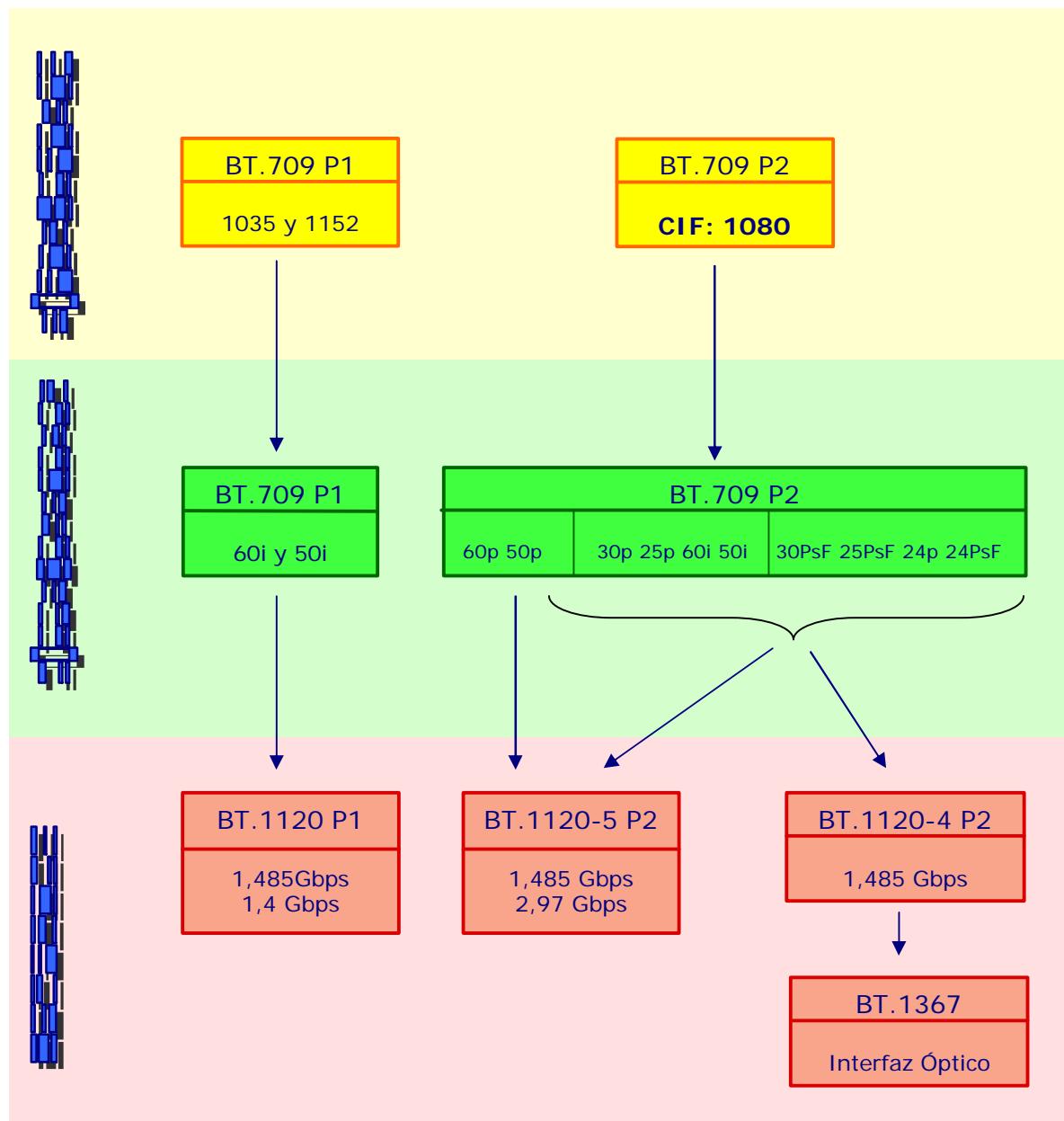


Figura 1. Normativa del UIT-R relativa a TVAD

³ Material Exchange Format

En su "Parte 1" la BT.709 hace referencia al sistema de TVAD analógico de primera generación, ya en desuso, y en su "Parte 2" es la utilizada por la totalidad de los actuales sistemas de TVAD digitales.

El interfaz HD-SDI se describe en la Recomendación **UIT-R BT.1120-4 P2** (Parte 2) cuya misión es la de encapsular los distintos formatos de muestreo definidos por la BT.709, así como los datos auxiliares más usuales: el audio digital embebido, los códigos de tiempos longitudinales y de intervalo vertical, los metadatos, identificadores del formato de vídeo o incluso vídeo comprimido en formato HD-SDTI. Su versión óptica es recogida en la BT.1367 [9].

Al igual que con la Recomendación BT.709, la Recomendación BT.1120-4 Parte 1 se corresponde a la anterior generación de TVAD, a pesar de que en la actualidad no existen equipos en el mercado con dicho interfaz.

2.3 Normativa SMPTE EN TVAD

La normativa del SMPTE (*Society of Motion Picture and Television Engineers*) referida a la TV de Alta Definición se apoya en el estándar colorimétrico del BT.709 Parte 2 del UIT-R, y tan solo aprobó para el antiguo formato de 1035 líneas visibles definido por la SMPTE 260M [10], una norma equivalente a esta de modo temporal e incompleta, recogida en el SMPTE 240M [11].

El SMPTE especifica 13 formatos distintos de producción, agrupados en las normas **SMPTE 295M** [12] y **SMPTE 274M** [13], equivalentes a la BT.709 P2; la **SMPTE 296M** [14] que define una resolución de 720 líneas visibles en formato progresivo, y la citada SMPTE 260M en desuso.

En la siguiente tabla se resumen los distintos parámetros de cada una de estas normas soportadas por el SMPTE para la actual TV de Alta Definición de 2^a Generación:

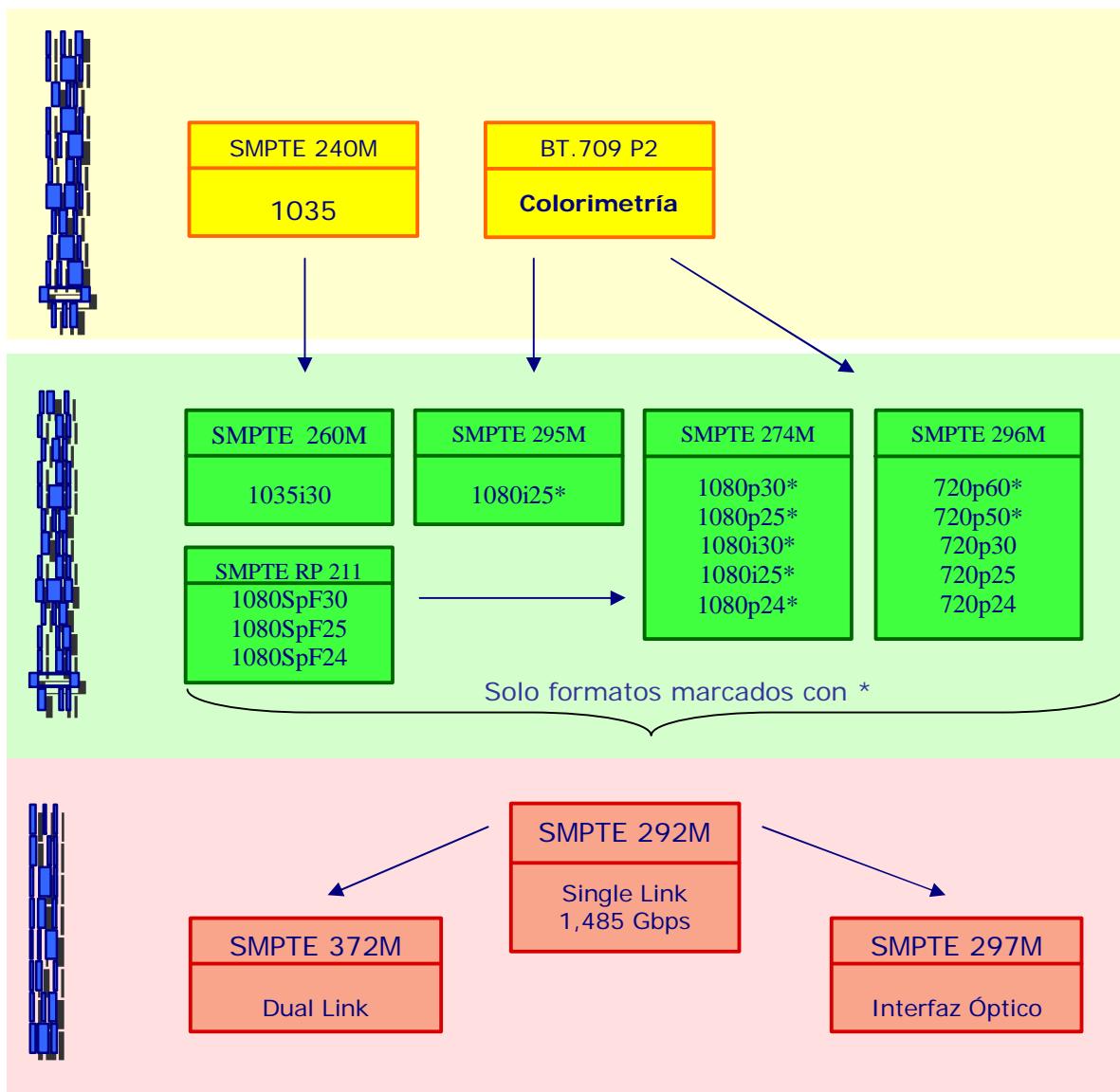
SMPTE	260M			295M			274M						296M		
	A	B	C	D	E	F	G	H	I	J	K	L	M		
Líneas Totales	1125	1125	1250	1125	1125	1125	1125	1125	1125	1125	1125	750	750		
Líneas Activas	1035	1035	1080	1080	1080	1080	1080	1080	1080	1080	1080	720	720		
Píxeles Activos	1920	1920	1920	1920	1920	1920	1920	1920	1920	1920	1920	1280	1280		
Frame Rate Scan	30	29,97	25	30	29,97	25	30	29,97	25	24	23,97	60	59,94		
Régimen Binario	2:1	2:1	2:1	2:1	2:1	2:1	1:1	1:1	1:1	1:1	1:1	1:1	1:1		
	1,485	1,483	1,485	1,485	1,4835	1,485	1,485	1,4835	1,485	1,485	1,4836	1,485	1,4835		

Tabla 1. Formatos de producción en TVAD definidos por el SMPTE

La convergencia con la industria cinematográfica facilitó la definición de los formatos de cuadro progresivos segmentados (PsF, *Progressive Segmented Frame*) con frecuencias de cuadro de 24Hz, 25Hz y 30Hz, recogidos en la recomendación **SMPTE RP 211** [15], y que no habían sido incluidos inicialmente en la SMPTE 274M. Utilizando el formato CIF, se puede transferir material cinematográfico de 1920x1080 con captura progresiva.

El interfaz eléctrico HD-SDI se define por el SMPTE en su recomendación **SMPTE 292M** [16], con total compatibilidad eléctrica y funcional con el especificado por el UIT-R en su BT.1120-4. Del mismo modo, se han aprobado normativas referentes al interfaz HD-SDI en modo Dual-Link para el transporte de formatos con submuestreos 4:4:4 y 4:4:4:4,

como SMPTE 372M [17]. Por último el interfaz óptico equivalente al BT.1367, queda definido en la norma SMPTE 297M [18].



Estándares y formatos de producción de contenidos

Figura 2. Normativa del SMPTE relativa a TVAD

En lo referente a los formatos de 1080p50 y 1080p60, en el 2006 el SMPTE llevó a cabo la estandarización de un interfaz eléctrico sobre cable coaxial en Single Link con velocidades de 3Gbps definido en la **SMPTE 424M**, así como el mapeo de los distintos formatos descritos sobre dicho interfaz en la recomendación **SMPTE 425M** [19].

Una vez más la convergencia con la industria cinematográfica en el denominado D-Cinema, ha llevado en el 2007 a la aprobación de la recomendación **SMPTE 435** [20] que define un interfaces digitales series de 10Gbps basado en fibra óptica. El D-Cinema trabaja con formatos de 2K y 4K de mayor resolución a los definidos en la actual TVAD, que exigen velocidades de 5.94Gbps y 8.91Gbps, muy superiores incluso a las requeridas por el formato 1080p50.

Recientes Test han demostrado la viabilidad de transmitir tasas binarias de 10Gbps utilizando modulaciones 4-level PAM sobre cable coaxial (Belden 1694A) con distancias próximas a los 100m, por lo que es previsible su futura estandarización y comercialización.

2.4 Recomendaciones EBU en TVAD

La EBU (*European Broadcasting Union*), como organización de reconocido prestigio en la elaboración de recomendaciones técnicas en el ámbito profesional de Televisión, ha tenido un papel muy activo y destacado en el análisis, definición y evaluación de los nuevos formatos de producción de TV en Alta Definición.

La EBU en su documento **EBU TECH 3299** [21], recomienda para el entorno de países que trabajan con resoluciones temporales de 50Hz, que se restrinja el amplio abanico de formatos estandarizados en Alta Definición a tan solo 4 formatos, 3 para la actual TVAD de 2^a Generación y uno para la futura 3^a Generación, con el objetivo de facilitar el intercambio e interoperabilidad entre radiodifusores.

En sus documentos técnicos, la EBU utiliza siempre para la definición de los distintos formatos de TVAD una nomenclatura basada en 3 campos (Número de líneas activas, barrido y cuadros por segundo), que difiere de la nomenclatura utilizada en numerosos documentos comerciales que utilizan como tercer campo el número de campos/s.

Los 4 formatos propuestos por la EBU para su implantación como formatos de intercambio de contenidos de TV en alta Definición son los siguientes:

Formato	Nomenclatura	Líneas activas	Píxeles/línea	Frame rate	Referencia SMPTE
System 1	720p50	720	1280	50 progresivos	296M
System 2	1080i25	1080	1920	50 entrelazados	274M
System 3	1080p25	1080	1920	25 progresivos	274M
System 4	1080p50	1080	1920	50 progresivos	274M

Tabla 2. Formatos de producción en TVAD definidos por la EBU

En lo concerniente al interfaz físico, la misma recomendación TECH 3299 establece como normativos la SMPTE 292M para medios físicos basados en coaxial y la SMPTE 297M para medios basados en Fibra Óptica.

La EBU en su documento **Statement D97** [22] de Enero del 2005, defiende como formato óptimo para la producción de contenidos el formato definido como "System 1" debido a su naturaleza progresiva, coincidente con la tecnología progresiva de las cámaras y los dispositivos de visualización no-CRT. Aun así, reconoce las dificultades existentes para adoptar dicho formato al tener escasa implantación comercial en su versión de 50Hz.

En la Recomendación **EBU R115** [23] de mayo del 2005, este organismo refuerza la necesidad de adoptar el formato 1080p50 (System 4) para la nueva generación de TVAD, avalado por la mejor calidad y eficiencia de los formatos progresivos frente a los formatos entrelazados, al considerar estos últimos como formatos heredados del entorno analógico de la década de los 50.

De igual modo en su recomendación **EBU TECH R112** [24], la EBU defiende la difusión de contenidos en formatos progresivos, apoyándose en distintas investigaciones científicas llevadas a cabo [25] [26] que demuestran la mejor calidad subjetiva ofrecida por el formato 720p50 frente a los formatos 1080i25 y 1080p25.

2.5 Producción de audio multicanal

2

Los radiodifusores que están introduciendo la TVAD en Europa consideran el sonido envolvente 5.1 como una parte fundamental e inseparable de la propia Alta Definición. En cualquier caso, implementar sonido envolvente conlleva una serie de desafíos como se describe en este apartado.

Se deben considerar 2 tipos de producciones de audio envolvente multicanal, en "Directo" y en "Postproducción". Las técnicas para crear contenido en directo en 5.1 están empezando a desarrollarse: los radiodifusores están produciendo en Alta Definición con sonido 5.1 en espacios deportivos y grandes eventos como entregas de premios y conciertos. Igualmente, las reglas de grabación y mezcla de las pistas de audio con sonido envolvente 5.1 para postproducción en géneros como el drama, han sido rápidamente establecidas gracias a la adaptación de las técnicas y herramientas del mundo del cine, a las especificaciones de tiempo y presupuesto de la televisión. Muchas de las grandes producciones ya son mezcladas con matriz surround para televisión e incluso en 5.1 para la versión en DVD.

Directo en 5.1

Para capturar eventos en directo, tales como en estadios y conciertos, son necesarios algunos micrófonos adicionales para hacer que el espectador se sienta transportado a las gradas o mezclado entre el público. Al igual que con las producciones en estéreo, micrófonos estereofónicos dirigidos al campo de juego o al escenario son normalmente usados para una perspectiva frontal con el comentario en el canal central. Micrófonos adicionales situados sobre el público en las gradas a la izquierda y derecha de la posición principal de las cámaras puede ser balanceado a los dos canales traseros (surround channels) para situar el murmullo de la grada o patio de butacas alrededor del telespectador.

Para programas como conciertos de música clásica, una imagen realista de la orquesta puede ser capturada usando técnicas estándares, como por ejemplo usar un Decca Tree⁴ posicionado sobre el director de orquesta para los canales izquierda, derecha y centro, donde el altavoz central añade estabilidad comparado con estéreo, y usar un Hamasaki Square⁵ (variante del IRT Cross⁶) sobre el público, capturando la reverberación del concierto en las paredes que puede ser reproducida en los canales traseros (Left surround, Right sound).

Una de las consideraciones más importantes cuando se produce material 5.1, es el ajuste de la metadata del audio. La Metadata es la información adicional relativa al audio, que es transportada junto con la señal de audio con la finalidad de permitir presentar el sonido en la mejor forma posible en el hogar. Por ejemplo, la metadata controla el downmix de 5.1 a estéreo para la reproducción de un programa de audio 5.1 en un aparato con sólo dos altavoces y, además, controla las características de la compresión del rango dinámico para altavoces pequeños, o para el modo "late night"⁷. La metadata permite transmitir la mejor mezcla posible con un rango dinámico sin compresión; y que sea el decodificador el que modifique estos canales de audio para que se adapte a las posibilidades del equipo de reproducción, ya sea éste tanto un equipo 5.1 como un televisor con un pequeño altavoz mono.

⁴ Array de micrófonos utilizado comúnmente en la grabación de orquestas

⁵ Array de 4 micrófonos bidireccionales utilizado en grabaciones con sonido envolvente

⁶ Array de 4 micrófonos cardioideos en cruz utilizado para la captación de sonido envolvente

⁷ Compresión del rango dinámico de la señal de audio para su adecuación a bajos niveles de escucha

Postproducción en 5.1

Para la postproducción de contenidos como documentales y dramas, crear sonido envolvente 5.1 no tiene por qué requerir más esfuerzo que para una mezcla estéreo. En producción lo más apropiado es grabar de forma separada tantos elementos como sea posible para dar la flexibilidad al posicionar las voces y al sincronizar cualquier efecto durante la mezcla.

En cualquier caso, la casi totalidad del diálogo reside en el canal central-frontal, de manera que la mayor parte de la dimensión en la imagen del sonido viene de la música y de los efectos de sonido que se añaden en postproducción. El diálogo es ocasionalmente balanceado, un poco fuera del centro para reflejar posiciones en la pantalla (esto es parecido a la mezcla en estéreo) con el altavoz central creando una posición más estable para los telespectadores sentados off-axis.

Hay que tener precaución y asegurar que el posicionado es efectivo en 5.1 y en estéreo (esto puede ser comprobado durante la mezcla usando una herramienta de monitorizado apropiada). La música en estéreo puede ser usada en la mezcla 5.1, aunque mezclas expresamente creadas para 5.1 dan un efecto mejor. Diferentes pistas con atmósferas estéreo pueden ser usadas en los altavoces frontales y traseros con el fin de sumergir al espectador en la acción.

2.6 Conclusiones

Como conclusiones en lo referido a la recomendación de los formatos de producción de contenidos de TV en Alta Definición, podemos exponer las siguientes recomendaciones:

- ? La industria profesional considera a la Alta Definición como el nuevo estándar de producción para televisión, ofreciendo una oferta que cubre todos los segmentos del mercado, desde la captación, postproducción, almacenamiento, y archivo.
- ? A diferencia de lo ocurrido en los entornos de definición estándar, la producción en Alta Definición presenta un escenario multiformato, coexistiendo en la actualidad 2 formatos progresivos denominados 720p50 y 1080p25, y un formato entrelazado denominado 1080i25.
- ? La recomendación de formatos progresivos desde distintas estancias, en particular desde la EBU, ha provocado la normalización de un nuevo formato de mayores prestaciones, denominado 1080p50, que se estima se convierta en el formato de producción a medio plazo.
- ? Dichos formatos están perfectamente regulados por los organismos competentes para ello, UIT y SMPTE, lo cual asegura una completa interoperabilidad entre ellos.

3 FORMATOS DE ALMACENAMIENTO E INTERCAMBIO DE PROGRAMAS

3

A finales de los años 90 surgieron tímidamente algunos formatos comerciales que facilitaban el almacenamiento de contenidos en Alta Definición utilizando para ello las tecnologías disponibles de compresión de vídeo, basados en sus inicios en soporte de cinta magnética, y que han ido evolucionando de forma vertiginosa en los últimos años, existiendo en la actualidad una amplia y variada oferta de formatos. Estos formatos cubren la totalidad de soportes físicos disponibles, cinta, disco óptico, memoria de estado sólido y disco duro, con tecnologías y prestaciones muy diversas.

La transición a la TVAD supone una fuerte ruptura con el actual escenario de definición estándar, donde prácticamente solo coexisten 2 formatos basados en las tecnologías MPEG-2 y DV con arquitectura Intracuadro, y con prestaciones muy similares (calidad, robustez, tasa binaria). Por el contrario, en TVAD la oferta de formatos incluyen una gran variedad de soluciones basadas en MPEG-2, MPEG-4 AVC, Wavelets, así como otros sistemas propietarios de elevadas prestaciones, que incorporan tanto arquitecturas Intercuadro como Intracuadro.

3.1 *Formatos de almacenamiento e intercambio de vídeo*

Los actuales formatos de almacenamiento en TVAD en el entorno de producción profesional, utilizan distintas tecnologías de compresión para reducir la tasa binaria de la fuente, y que pueden ser caracterizados por los siguientes parámetros:

1. Profundidad de píxel: al igual que en campo de la definición estándar, la mayoría de los formatos de TVAD realizan su compresión inicialmente con una precisión de **8 bits/muestra**, pero han aparecido nuevos formatos para producciones de mayor calidad que facilitan el procesamiento con precisión de **10 bits/muestra**.
2. Resolución espacial: a pesar de que inicialmente solo se comercializaron productos para los formatos **1080i25** y **1080p25**, en la actualidad la práctica totalidad de los nuevos productos soportan también el formato **720p50**. Es importante constatar que a día de hoy ya existe una solución comercial que dan soporte de almacenamiento al formato 1080p50 (HDCAM-SR).
3. Submuestreo de la componente de luminancia: existen formatos que a pesar de admitir en sus interfaces de entrada y salida los formatos estándar de 1080 y 720 líneas, llevan a cabo ligeras compresiones espaciales en el dominio horizontal (**1440 píxel/línea** y **960 píxel/línea**) con el objetivo de reducir la tasa binaria necesaria para su codificación.
4. Submuestreo de las componentes de color: el submuestreo horizontal y en ocasiones vertical de las componentes de crominancia, conlleva la definición de los formatos denominados **4:4:4**, **4:2:2**, **4:2:0** y **3:1:1**.
5. Modelo Predictivo Temporal: es conocido que la codificación Intracuadro facilita las tareas de edición y postproducción, así como los requerimientos computacionales de los sistemas, pero también requieren de una tasa binaria sustancialmente superior a la ofrecida por las arquitecturas Intercuadro. Es por ello que en la actualidad coexisten ambos formatos **Intracuadro e Intercuadro** con GOPs de 12-15 cuadros.

6. Tasa binaria: las tasas binarias requeridas en TVAD ofrecidas por los distintos formatos comerciales cubren un amplio espectro: 50Mbps, 75Mbps, 100Mbps, 140Mbps, 185Mbps, 440Mbps, e incluso 880Mbps para aplicaciones de extrema calidad con submuestreo 4:4:4.
7. Tecnología de codificación: En la TVAD aparecen nuevos formatos de codificación alternativos al clásico **MPEG-2**, como son el **MPEG-4 AVC/H.264** en sus perfiles profesionales y **JPEG2000**. A pesar de la teórica mejora en la eficiencia de estos formatos con respecto a la obtenida por MPEG-2, es cierto que su grado de madurez y optimización dista de la alcanzada actualmente por MPEG-2.
8. Soportes de almacenamiento: los formatos de almacenamiento físico incluyen cinta magnética de $\frac{1}{2}$ ", disco óptico basado en el "*Professional Disc*", memoria de estado sólido "*Compact Flash*" y P2, disco duro "*FieldPack2*", así como las nuevas ExpressCard que sustituyen a las actuales PCMCIA. Por lo general todos los formatos admiten la encapsulación de sus archivos en formato MXF que permite independizar el contenido del soporte físico, facilitando su intercambio.

En los siguientes subapartados se describen las soluciones tecnológicas adoptadas por las distintas alternativas comerciales, en función de la arquitectura de compresión utilizada.

3.1.1 MPEG-2 Intercuadro

El formato **XDCAM HD fase III** ofrece una solución profesional en TVAD para grabación sobre disco óptico "*Professional Disc*", basado en compresión MPEG-2 [27] en su configuración 422P@HL en modo Intercuadro, lo que le permite trabajar con una tasa binaria de tan solo **50Mbps** al explotar ambas correlaciones, temporal y espacial. Para mantener su calidad en las etapas de edición, incorpora el modo "*Smart Rendering*" que le permite mantener el correcto alineamiento del GOP en todo el proceso.

Este formato no aplica ningún tipo de submuestreo a las componentes de luminancia y crominancia, por lo que permite trabajar en resolución completa en cualquiera de los formatos 1080 y 720. Para el archivo de contenidos en ficheros informáticos utiliza encapsulación universal MXF (OP-1a), facilitando la interoperabilidad con los sistemas de edición y postproducción basados en infraestructuras IT.

3.1.2 AVC-Intra

El formato **AVC-Intra Class 100** está basado en los nuevos perfiles profesionales definidos en el estándar MPEG-4 AVC/H.264. En Mayo del 2003, dicho estándar, al igual que lo ocurrido con MPEG-2 en 1993, fue aprobado de modo conjunto por la UIT, denominándolo H.264, y por el ISO/IEC denominándolo ISO/IEC 14496-10, y conocido como MPEG-4 AVC.

A pesar de que inicialmente solo se definieron 3 perfiles (**Baseline**, **Extended** y **Main**) optimizados para aplicaciones de difusión adaptados a diferentes condiciones de retardo, complejidad y protección de errores; su elevada eficiencia de codificación provocó en julio del 2004 la aprobación de 4 nuevos perfiles adaptados a los requerimientos exigidos en los entornos profesionales de producción.

Dichos perfiles nacen como una evolución del "Main Profile" al que incorporan todos ellos 3 nuevas herramientas: la Transformada Adptativa 8x8 vs 4x4, la cuantificación perceptual y la cuantificación independiente para las componentes de crominancia; diferenciándose tan solo en dos aspectos: la profundidad de píxel y el tipo de submuestreo de la información de crominancia, como se muestra en la siguiente tabla:

Perfil	Denominación	Submuestreo o Crominancia	Profundidad de Píxel (Máxima)
High	HP	4: 2: 0	8 bits/muestra
High 10	Hi10P	4: 2: 0	10 bits/muestra
High 4:2:2	Hi422P	4: 2: 2	10 bits/muestra
High 4:4:4 Predictive	Hi444P	4: 4: 4	14 bits/muestra

Tabla 3. Características de los Perfiles definidos en MPEG-4 AVC/H.264

En abril

del 2007, fueron incorporados 4 nuevos perfiles específicos para los entornos profesionales de producción al utilizar una arquitectura Intracuadro. Estos perfiles son los denominados High 10 Intra, High 4:2:2 Intra, High 4:4:4 Intra y CAVLC 4:4:4 intra.

El formato AVC-I Class 100, está basado en el Perfil **High 422 Intra**, con submuestreo 4:2:2, 10 bits/muestra y una tasa binaria de **100Mbps**. Al igual que ocurre con el formato XDCAM HD fase III, este formato no aplica ningún tipo de submuestreo a las componentes de luminancia y crominancia, trabajando a resolución completa en cualquiera de los formatos 1080 y 720 soportados.

Como soporte físico utiliza memorias de estado sólido P2, utilizando encapsulación universal MXF (OP-Atom y OP-1a) conforme a la SMPTE RP 2008 [28].

3.1.3 JPEG2000

Los nuevos formatos **Infinity** están basados en el formato conocido como JPEG2000 [29], y estandarizado por el ISO (International Organization for Standardization) en el año 2000 como ISO/IEC 15441-1. Este estándar se diferencia básicamente de los anteriores (MPEG-2 y MPEG-4 AVC/H.264) en la utilización de una codificación transformacional global basada en wavelets, en sustitución de la tradicional transformada DCT, que opera de modo local en áreas no solapadas de reducidas dimensiones.

JPEG2000 nació con la intención de sustituir al estándar JPEG en los entornos informáticos, ya que ofrece una eficiencia de compresión hasta 30 veces superior a este, pero sorprendentemente su primer campo de aplicación ha sido en el entorno profesional de alta calidad, al ser el formato adoptado por la industria cinematográfica para la distribución de Cine Digital.

La familia Infinity HD trabaja en modo Intracuadro, con resoluciones nativas de 1080 y 720 (sin ningún tipo de submuestreo), con 10 bits/muestra, formato 4:2:2 y tasas binarias de **50Mbps**, **75Mbps** y **100Mbps**. Como soporte de almacenamiento físico admite soporte de disco duro “REV PRO drive” y memorias de estado sólido “*Compact Flash*”, con encapsulado MXF (OP-1a) en ambos casos.

3.1.4 Otros formatos alternativos

Entre los formatos profesionales alternativos cabe destacar la familia **HDCAM** y **HDCAM-SR**, utilizada en la ultima década en innumerables producciones de indudable prestigio, utilizando como soporte físico la cinta magnética de 1/2”. La familia de formatos **HDCP PRO HD** permite el almacenamiento en cinta magnética de 1/4” y en

memorias de estado sólido P2, y ha tenido, de igual modo, una gran implantación en la producción de contenido en TVAD.

En lo referente a estaciones de edición y postproducción, existen soluciones profesionales que integran codecs propietarios como el DNxHD y ProRes 422, y que gracias a sus prestaciones y a la elevada tasa binaria con la que trabajan, próxima a los 200Mbps, ofrecen una gran robustez, soportando fuertes procesos de postproducción sin perdida apreciable de calidad.

3.2 Formatos de almacenamiento e intercambio de audio multicanal

Las producciones en audio multicanal, una vez creadas junto con su metadata asociada, deben ser almacenadas y distribuidas en la red del radiodifusor para su archivo e intercambio. Para ello existen en la actualidad distintas alternativas:

- ☛ Procesar las distintas señales como canales de **audio discretos PCM**, embebiendo dichos canales en la trama HD-SDI y embebiendo su metadata en el VANC. Esta solución presenta como ventaja su bajo coste y alta calidad, al no aplicar ningún tipo de codificación a los canales de audio, pero requiere de un tratamiento transparente de todos los equipos de la cadena de producción para evitar la perdida o corrupción de la metadata, condición de difícil validación en muchos casos. De igual modo, demanda el poder almacenar 6 pistas de audio en los grabadores y dispositivos de almacenamiento, circunstancia no siempre disponible en numerosos equipos comerciales.
- ☛ El método más extendido consiste en utilizar un codificador **Dolby-E**. Este formato permite transportar hasta 8 canales discretos de audio y su metadata asociada, en un par AES/EBU para su almacenamiento en los distintos soportes físicos como cintas, discos ópticos, memorias o servidores. El formato Dolby-E es muy robusto, soportando hasta 5 multigeneraciones sin perdida apreciable de calidad. Su alineamiento de trama en intervalos de 40ms facilita la edición y corte sin clicks ni mutes. Tanto la facilidad de multigeneración como la de alineamiento de trama no son soportadas por otros codificadores multicanal optimizados para entornos de difusión.
- ☛ Existen otras alternativas de formatos de audio codificación multicanal con baja penetración comercial, como es la utilización de **DTS**, **APTX** y otras soluciones propietarias de distintos fabricantes, con características y funcionalidades similares a las descritas para Dolby-E.

La utilización de un determinado formato de almacenamiento para audio multicanal debe cumplir con los siguientes requisitos:

1. Preservación de la metadata creada en la fase de producción
2. Tecnología de codificación eficiente, con el objetivo de ofrecer una tasa binaria reducida sin comprometer la calidad de la señal.
3. Fuerte robustez a la multigeneración y trascodificación.
4. Capacidad de llevar a cabo tareas de postproducción sencillas (corte) sin necesidad de decodificar toda la trama.

5. Absoluta interoperabilidad con los distintos codificadores multicanal de difusión como AC3 y HE-AAC.
6. Bajo retardo de codificación que no genere desfases audio-vídeo difíciles de compensar.
7. Capacidad de encapsulado en archivos informáticos como MXF y Broadcast Wave.

La UIT en su recomendación **BS.1548-1** [30], expone en su Anexo I los principales requisitos (calidad, ancho de banda, retardo, concatenación de codecs, bit rates) para los codificadores de audio multicanal en los entornos de producción y contribución, requisitos reunidos tan solo por el formato Dolby-E.

3.3 Recomendaciones

Como conclusiones en lo referido a la recomendación de los formatos más apropiados para almacenamiento e intercambio de contenidos de TV en Alta Definición, podemos exponer las siguientes recomendaciones:

- El conjunto de formatos presentan una madurez suficiente para entender que existirá un entorno multiformato con diversidad de soportes físicos.
- Las producciones en Alta Definición son entendidas mundialmente como un servicio de alta calidad, y es por ello que se debe preservar al máximo la calidad en los procesos de producción e intercambio de contenidos, utilizando formatos de altas prestaciones.
- La utilización del formato 1080p50 se encuentra en una fase inicial, existiendo ya dispositivos comerciales de conmutación, distribución, captación y almacenamiento, por lo que es previsible que en los próximos años se amplíe dicha oferta.
- Se recomienda la utilización de formatos 4:4:4 para producciones con calidad cinematográfica donde se requieran una elevada resolución colorimétrica. De igual modo, se recomienda la utilización del formato 4:2:2 para producciones de alta calidad donde se llevan a cabo fuertes elaboraciones de postproducción. Existe un amplio entorno de aplicación como producción de informativos y reportajes que permiten la utilización de formatos 4:2:0.
- Se recomienda la utilización de formatos de audio multicanal para el almacenamiento y distribución de contenidos de audio multicanal, que ofrezcan una elevada y robustez a los procesos de multigeneración, sin pérdida perceptual de calidad subjetiva.

El proceso de contribución de contenidos de TV digital comprende la etapa de transmisión de las señales audiovisuales desde los puntos en los que se captan (eventos deportivos, noticias,...) hasta los centros de producción de los radiodifusores para su posterior modificación y procesado, así como al intercambio de señales entre estudios o entre radiodifusores. Este proceso siempre es previo a la emisión final del contenido.

Los contenidos recibidos en los centros de producción son grabados (recodificados) y editados una o varias veces, donde por lo general son insertados logos o grafismo antes de ser emitidos, y en definitiva, sufrirán una serie de manipulaciones que requieren que la fuente de la señal sea de muy alta calidad, para que el contenido que finalmente se emite no sufra una degradación excesiva de la calidad.

Estos contenidos serán transportados por las redes de telecomunicaciones habituales (fibra óptica, satélite, redes SDH/PDH o redes IP), que a día de hoy permite la transmisión de hasta 16 señales HD-SDI sin compresión por fibra óptica utilizando tecnología CWDM, incluso 64 señales de Alta Definición si se aplica una ligera compresión a cada una de las fuentes.

Si se desea incurrir en menores costes por parte de los radiodifusores, se deberá aplicar una mayor compresión para su adaptación a redes de menor capacidad. Estos compromisos entre alta calidad, compresión y coste llevan a definir los formatos de contribución para señales de TV de Alta Definición, que a día de hoy podemos resumir en los tres principales formatos de codificación de vídeo MPEG-2, MPEG-4 AVC/H.264 y JPEG2000.

4.1 MPEG-2

El formato MPEG-2 es el formato más usado y extendido para la codificación de TV en definición estándar. En lo referente a contribución, el formato MPEG-2 definió el "**Perfil 4:2:2**", que utiliza el modelo de muestreo de las señales de luminancia y crominancia compatible con los equipos de postproducción de TV digital existentes, evitando la degradación de la calidad de la señal que se produciría con los modelos de muestreo 4:2:0.

Para la codificación de señales de TV de Alta Definición, MPEG-2 definió el "**High Level**", que soporta resoluciones de hasta 1920x1152 como máximo, con una definición de 8 bits por muestra y bit rates de compresión hasta 100 Mbps como máximo.

Por lo tanto, el formato MPEG-2 que se utiliza para la contribución de señales de TV en Alta Definición es el **422P@HL**. Este es el formato que se ha venido utilizando desde la implantación de la TV de Alta Definición para la contribución de contenidos como estándar de "facto".

Se trata de un formato maduro, que implementan la mayor parte de los codificadores desarrollados por los fabricantes para aplicaciones de contribución de TV de Alta Definición desde hace unos 5 años, y que introduce un bajo retardo en la transmisión debido al proceso de codificación y decodificación (0,5 segundos), razonable para aplicaciones de contribución en directos.

En cuanto al bit rate de compresión utilizado por este formato, se utilizan valores entre 3 y 4 veces superiores a los valores usado por la TV de definición estándar, por lo que los valores típicos usados por los radiodifusores suelen variar entre 30 y 60 Mbps.

El formato MPEG-2 422P@HL ha sido el formato utilizado para la contribución de los principales eventos producidos Alta Definición hasta la fecha (Campeonato del Mundo de

fútbol en Alemania 2006, Champions League, Olimpiadas de invierno en Torino 2006) y va a ser el formato utilizado como formato de contribución por la mayoría de los radiodifusores en los Juegos Olímpicos de Beijing en el año 2008.

4.2 MPEG-4 AVC/H.264

El formato de compresión MPEG-4 AVC/H.264 fue desarrollado con el objetivo de mejorar las prestaciones del formato MPEG-2, sobre todo a bit rates bajos. Al igual que en MPEG-2, en MPEG-4 también se han definido un "**Perfil 422**" adecuado a los requerimientos exigidos en los entornos de producción. El Nivel requerido para los formatos 720p50, 1080i25 y 1080p25 es el "**Nivel 4**".

Para aplicaciones de contribución, la UIT-R en su recomendación BT.1737 [31] especifica la utilización en TVAD del formato 422P@L4 y el formato Hi10P@L4, este último tan solo para servicios de SNG. Pero a día de hoy no existe ningún fabricante que implemente estos perfiles, y los planes de desarrollo e implementación de estos formatos se demorarán hasta finales de 2008, debido a la falta de chipsets.

Por lo tanto, si se desea utilizar el formato de compresión MPEG-4 AVC para aplicaciones de contribución, a día de hoy, se deberá utilizar el mismo Nivel y Perfil que se utiliza para las aplicaciones de distribución y difusión, este es el HP@L4, a un bit rate máximo de codificación de 25 Mbps, una resolución de 8 bits por muestra y con muestreo 4:2:0.

Son varios los motivos que desaconsejan la utilización de este formato para aplicaciones de contribución:

- Muestreo de las señales de luminancia y crominancia 4:2:0, con la consiguiente pérdida de calidad en los procesos de postproducción y recodificación posterior de los contenidos.
- La poca eficiencia del formato de compresión MPEG-4 AVC a elevados bit rates, que hace que el ahorro de ancho de banda con respecto a MPEG-2 no sea tan importante como a bit rates bajos.
- El actual retardo que introducen los codificadores MPEG-4 AVC, debido a la complejidad de los algoritmos, puede llegar a estar entre 3 y 4 segundos, lo que para muchas aplicaciones de contribución (eventos deportivos o en vivo) supone una limitación muy importante.

4.3 JPEG2000

El formato de compresión JPEG2000 ha sido desarrollado para la compresión de vídeo de muy alta calidad, por lo que precisa de redes de telecomunicaciones de elevado ancho de banda para su utilización. Es el formato de compresión que ha definido la norma DCI⁸ para la codificación y distribución de cine digital, en donde se habla de resoluciones superiores a las definidas para la TV de Alta Definición de 2K y 4K.

Los avances en cuanto a capacidad, disponibilidad y calidad de redes de telecomunicaciones (redes IP con capacidad Gigabit y QoS) hace que a día de hoy este formato de compresión se esté comenzando a utilizar por parte de algunos radiodifusores para servicios de contribución de TV de Alta Definición.

JPEG2000 presenta algunas ventajas respecto a los formatos MPEG-2 y MPEG-4 AVC:

- Es un algoritmo de codificación basado en wavelets en lugar de la DCT (*Discrete Cosine Transform*). Esta transformada se aplica sobre imágenes completas en

⁸ Digital Cinema Initiative

lugar de aplicarse sobre bloques de 8x8 o 16x16, evitando el efecto de bloques típico de la codificación MPEG.

- Es un formato de codificación desarrollado y optimizado para la compresión de vídeo a alta calidad, al contrario que MPEG-2 y MPEG-4 que fueron desarrollados y optimizados para la compresión a bajo bit rate. El tipo de muestreo utilizado es 4:2:2 y la resolución es de 10 bits por muestra.
- El retardo introducido por el proceso de codificación y descodificación en JPEG2000 es muy bajo, en torno a 100-150 milisegundos, debido a que no se utilizan dependencias entre imágenes.

Como principal desventaja de este formato de codificación está el elevado ancho de banda requerido, que oscila entre los 50 y 200 Mbps (típico 100–120 Mbps), siendo necesarias redes IP privadas con QoS o enlaces dedicados STM-1 para su utilización.

JPEG2000 ha sido seleccionado por el operador sueco Telenor para la contribución de los partidos de fútbol de la Premier League inglesa en Alta Definición desde Londres a Suecia utilizando redes IP con capacidad Gigabit Ethernet.

También ha sido el formato utilizado por la EBU en las demostraciones realizadas durante el IBC-2007 para el intercambio de contenidos de TVAD en entornos de producción. Las tasas de compresión utilizadas por la EBU en estas demostraciones fueron de 115 Mbps (1080i/25 y 720p/50) y 200 Mbps (1080p/50).

4.4 Recomendaciones

Como conclusiones en lo referido a la recomendación de los formatos de compresión de vídeo más apropiados para la contribución de TV de Alta Definición, podemos exponer las siguientes recomendaciones:

- El formato recomendado a día de hoy en cuanto a madurez, disponibilidad y calidad ofrecida es el MPEG-2 422P@HL. Este es el formato que se ha venido utilizando hasta la fecha y que se va a seguir utilizando de forma generalizada al menos durante los próximos dos años.
- La utilización del formato MPEG-4 AVC/H.264 para aplicaciones de contribución, no parece viable a día de hoy debido a la no disponibilidad de equipos comerciales que soporten el Perfil 422. Se estima que la tecnología para la utilización de este tipo de formatos estará madura a partir del año 2009.
- El formato JPEG2000, se posiciona como un formato muy interesante de cara a un futuro próximo, aunque debido al elevado ancho de banda requerido, supone una barrera difícil de superar.

5 FORMATOS DE DIFUSIÓN DE CONTENIDOS

El proceso de emisión de contenidos es sin duda el factor clave que determina en gran medida los dos aspectos más relevantes de la cadena de difusión, como son la calidad final percibida por el usuario y el ancho de banda requerido para su difusión. Esta etapa cobra una especial relevancia por el hecho de afectar de modo transversal al conjunto de medios de transmisión como son el Satélite, Cable, la Radiodifusión Terrestre o el IPTV.

Con objeto de alcanzar con la mayor celeridad una economía de escala en el mercado de consumo de la TVAD, sería deseable que el conjunto de plataformas utilizaran los mismos formatos de difusión o al menos formatos con un alto grado de interoperabilidad.

En este apartado se detallan las alternativas existentes a la codificación de vídeo tradicional basadas en compresión MPEG-2 y las soluciones tecnológicas existentes para la prestación de servicios de audio multicanal 5.1, que enriquecen en gran medida los contenidos más demandados en Alta Definición como son los programas deportivos, los conciertos y operas, o las películas.

Debido al amplio espectro de soluciones tecnológicas disponibles, la adopción de un determinado formato de emisión debe ser acorde en todo momento con las capacidades comerciales de los actuales receptores domésticos, con el objetivo de alcanzar una total interoperabilidad. Este hecho exige una especial coordinación entre radiodifusores, operadores y fabricantes de modo que se alcance un equilibrio entre calidad, costes y ancho de banda consumido.

5.1 Estándares de codificación de vídeo

Cuando el sistema de codificación MPEG-2 fue normalizado en 1993, representaba el estado más avanzado de la tecnología de codificación y en consecuencia fue adoptado por el DVB para todas sus normas de difusión. Estas normas se referían todas a servicios de definición estándar, porque el DVB no contemplaba en aquel momento la TVAD como servicio comercial real, probablemente como consecuencia del fracaso del formato Europeo HDMAC⁹ para Alta Definición. La codificación MPEG-2 de una señal de TVAD requería en aquel entonces una tasa binaria de entre 19-20 Mbps, es decir, la capacidad total disponible en los actuales multiplex de TDT en España.

La norma MPEG-2 define sólo el decodificador, permitiendo la aplicación de innovaciones tecnológicas al codificador. El resultado ha sido que en los años transcurridos desde 1993 el flujo medio para codificar una señal de TV de definición estándar ha disminuido en un 40-45% del flujo requerido en 1993. Pero su capacidad de mejora futura está prácticamente agotada. Por tanto, aplicado a la codificación de la TVAD, sin condiciones para la multiplexión estadística, es difícil esperar que MPEG-2 sea una solución económicamente viable para modelos de difusión con fuertes restricciones de ancho de banda, como son los entornos de IPTV, TDT y en menor medida satélite y/o cable HFC.

5.1.1 MPEG-4 AVC/H.264

MPEG-4 AVC/H.264 ha sido diseñado con el objetivo de mejorar la eficacia de MPEG-2, reduciendo el flujo necesario para la misma calidad de la señal vídeo en un 50%, o incluso menos del 50% para ciertas aplicaciones. Para ello incluye un grupo numeroso de tecnologías, alguna de ellas con un alto grado de complejidad, y que tienen una alta expectativa de mejora futura, lo que redundaría en una probable mejora continuada de la

⁹ High Definition Multiplexed Analog Components

eficacia de codificación de MPEG-4 AVC/H.264 a lo largo de su vida útil, que se espera supere ampliamente el porcentaje de mejora de eficiencia de codificación de MPEG-2¹⁰.

En particular la probable futura inclusión de desarrollos previstos, todavía no incorporados, como el SVC¹¹ puede incrementar de forma notable la eficacia de codificación y la versatilidad de la norma MPEG-4 AVC/H.264. Por estas consideraciones H.264 es un sistema recomendado para codificar programas de TVAD.

La difusión de contenidos en Alta Definición con MPEG-4 AVC/H.264, puede realizarse tanto con el "**Main Profile**" (MP), como con el "**High Profile**" (HP), pero hay que hacer notar que, siendo este estándar un sistema muy complejo, las primeras implementaciones de chipset de codificación han utilizado sólo las tecnologías básicas, y han ido añadiendo nuevas tecnologías de forma progresiva. Los chipset disponibles en este momento para el perfil HP son todavía relativamente simples, por lo que su eficacia de codificación es hoy un tanto inferior al objetivo del 50%, pero se espera que incorporen diversas facetas opcionales de la norma lo largo de 2008 para conseguir ese objetivo del 50% de eficacia.

Otra tecnología, independiente del sistema de codificación, pero íntimamente ligada al mismo, es la de los algoritmos de multiplexión estadística. Los algoritmos para multiplexión estadística de señales MPEG-4 AVC/H.264 están siendo perfeccionados y es probable que se aproximen a la eficacia conseguida en el caso de señales MPEG-2, aunque es muy difícil que lleguen a superarla, por razón de la menor redundancia residual de las señales MPEG-4 AVC/H.264 frente a las MPEG-2.

5.1.2 Otras tecnologías de codificación alternativas

En paralelo con el desarrollo de MPEG-4 AVC/H.264, Microsoft desarrolló otro sistema de codificación alternativo y pensado originalmente para aplicaciones de "streaming" en Internet. Estaba basado en Windows Media¹², sistema propietario de Microsoft, pero representaba sólo una versión reducida de este último, por cuanto no incluía facetas tales como DRM, metadata, interfaz de usuario, etc. Este sistema fue normalizado por el SMPTE en 2006, tras más de dos años de proceso, como **VC-1**[32]. VC-1 estaba diseñado para exploración progresiva, preferentemente para aplicaciones de flujo pequeño (inferior a 1 Mbit/s), pero es susceptible de ser utilizado en aplicaciones con mayor demanda de flujo. Es bastante menos complejo que MPEG-4 AVC/H.264¹³, pero también su eficacia de codificación es más pequeña.

Las demostraciones de TVAD organizadas por la EBU en la IBC 2005, orientadas a la comparación de formatos (1080i y 720p) a diferentes valores de flujo y con diferentes procedimientos de codificación, permitieron constatar que la calidad de imagen conseguida para un mismo valor de flujo era apreciablemente mayor con MPEG-4 AVC/H.264 que con una versión avanzada de VC-1.

Es probable que MPEG-4 AVC/H.264 y VC-1 coexistan en algunas aplicaciones, posiblemente en mercados verticales, y sea el mercado quien se decante por una u otra norma. En las normativas de DVB se ha incluido recientemente en sus especificaciones el

¹⁰ La mejora de eficacia de MPEG 2 ha permitido reducir el flujo necesario para la misma calidad de vídeo en un 10-15% por año, en los primeros años de su vida.

¹¹ SVC (scalable video coding),

¹² Windows Media 9 en el momento de la propuesta de normalización de VC-1

¹³ La menor complejidad de VC-1 podría significar hoy un coste menor de los chips correspondientes, pero en muy poco tiempo la economía de escala diluye las diferencias de costes. Es más, es muy probable que muchos chips, como los previstos para los DVD de Alta Definición, integren decodificadores de ambos sistemas.

uso de uno u otro sistema en sus normas aplicables a TVAD. Algo parecido ha sucedido con el formato Blu-Ray DVD, que permiten ambas normas. Sin embargo en Europa no hay referencias hasta ahora de la adopción de VC-1 para radiodifusión (TVAD, TV de pago, DVB-H, etc.). En América el estándar VC-1 tiene más aceptación, pero su adopción es limitada.

5.2 Estándares de codificación de audio multicanal

La gran aceptación en el mercado de consumo de los sistemas de reproducción de sonido multicanal envolvente, denominados "Home Cinemas" o "Home Theater", así como la incorporación de decodificadores de audio multicanal en las pantallas de nueva generación, permite a los radiodifusores enriquecer los servicios de TV de Alta Definición incorporando sonido envolvente 5.1 a sus producciones, que provoca una mejora sustancial de la sensación subjetiva percibida.

La producción de contenidos con sonido multicanal no se limita a los géneros musicales o films, sino que puede se pueden explotar igualmente en eventos en directo, como son los eventos deportivos, y que se ha demostrado generan en el espectador una increíble sensación de realismo. Como ejemplo más destacado, cabe reseñar que por primera vez la producción unos Juegos Olímpicos, en concreto los de Pekín, se van a producir en su totalidad con sonido multicanal 5.1.

En los siguientes apartados se describen los dos codificadores de audio AC3/E-AC3 y HE-AAC, recomendados para la difusión de servicios de audio multicanal, por distintos organismos internacionales como el EICTA, la EBU y el DVB.

5.2.1 AC3, E-AC3

Los codecs AC3 y E-AC3 conocido como DD y DD+¹⁴ respectivamente, son formatos propietarios Dolby Laboratorios, pero están recomendados por la mayoría de los estándares abiertos como el DVB, estandarizado como ETSI TS 102 366 [33], y el ATSC donde el codec AC3 figura como obligatorio para la difusión de servicios de TDT.

El codec AC3 dispone de capacidad para aplicaciones mono, estéreo con tasas binarias entre 128Kbp y 192Kbps, y multicanal hasta 5.1, cuya configuración de 448Kbps es la más extendida entre los radiodifusores. Pero la gran ventaja ofrecida por este codec reside en las posibilidades que ofrece y transmitir la metada que se genera en la fase de producción, entre la que cabe destacar:

- ☞ **Normalización de diálogos:** metadata que asegura un nivel de escucha consistente entre los distintos programas, evitando las variaciones bruscas en la commutación de servicios.
- ☞ **Downmix a estéreo:** metadata que permite generar en el receptor una versión estéreo de la señal multicanal recibida, evitando el tener que transmitir en simulcast dichos canales, con el consiguiente ahorro de ancho de banda que ello implica.
- ☞ **Control del rango dinámico:** metadata que asegura una reproducción óptima para cada uno de los diferentes ambientes de escucha.

¹⁴ Dolby Digital y Dolby Digital Plus

El nuevo esquema de codificación E-AC3¹⁵ mejora la eficiencia de su predecesor AC3, aumentando su capacidad hasta configuraciones 7.1, y manteniendo su retrocompatibilidad con éste, al transportar siempre una versión AC3 en su trama, facilitando su decodificación por los actuales dispositivos comerciales como "Home Cinemas".

Con el objetivo de reducir sustancialmente las tasas binarias requeridas, E-AC3 incorpora nuevas herramientas de codificación como el procesador de pre-noise, la codificación transformacional basada en DCT, la extensión espectral o la cuantificación con ganancia adaptativa, entre otras.

5.2.2 HE-AAC v2

El codificador de audio HE-AAC v2¹⁶, estandarizado como ISO/IEC 14496-3 [34] y conocido comercialmente como MPEG-4 aacPlus v2, permite la codificación de audio en versiones mono, estéreo o multicanal, con hasta 48 pistas.

A diferencia del codec AC3, el codec HE-AAC está optimizado para aplicaciones con fuertes restricciones de ancho de banda, motivando este hecho su obligatoriedad en estándares como el S-DMB, ISDB, 3GPP, y siendo estandarizado recientemente en la especificación técnica ETSI TS 102 005 [35] para la difusión de servicios DVB sobre entornos IP.

HE-AAC nace como una evolución del codec MPEG-4 AAC basado en los modelos de enmascaramiento psicoacústico tradicional, al que incorpora 2 nuevas tecnología:

- ☞ **Spectral Band Replication:** explota la alta correlación entre las bandas de bajas y altas frecuencias, que permite la reconstrucción de la señal con la información de la envolvente espectral "SBR data"¹⁷. Esta técnica es la utilizada en el codec conocido como aacPlus V1.
- ☞ **Parametric Stereo:** herramienta que explota las diferencias de intensidad (IID)¹⁸ y de fase (IPD)¹⁹ de la imagen estéreo de la señal de audio, así como su correlación cruzada (ICC)²⁰.

Una de sus principales ventajas de los decodificadores HE-AAC v2 es su retrocompatibilidad con los streams aacPlus v1 y AAC. Distintas evaluaciones de calidad [36] han verificado la alta eficiencia de este codec, ya que permite para la mayoría de los contenidos, las codificaciones de señales multicanal 5.1 utilización tasas binarias próximas a los **160Kbps**. Esto supone una reducción sustancial de ancho de banda, comparado incluso con la tasa binaria utilizada en la actualidad para la codificación de servicios de TV en estéreo (MPEG-1 Layer II con 192Kbps).

Como inconveniente, cabe reseñar que los receptores con capacidades de decodificación HE-AAC v2, requieren para una correcta interoperabilidad con los dispositivos de reproducción de dispositivos domésticos "Home Cinema", la recodificación de la señal multicanal decodificada, bien al formato DTS o a AC-3.

¹⁵ Enhanced AC3

¹⁶ High Efficiency Advanced Audio Coding Version 2

¹⁷ Spectral Band Replication data

¹⁸ Inter-Channel Intensity Difference

¹⁹ Inter-Channel Phase Difference

²⁰ Inter-Channel Cross-Correlation

5.3 Recomendaciones

Como conclusiones en lo referido a la recomendación de los formatos de difusión más apropiados para la emisión de servicios de TV de Alta Definición, podemos exponer las siguientes recomendaciones:

1. El conjunto de formatos y estándares utilizados para la difusión de contenidos de TV en Alta Definición, deben estar comprendidos entre los recogidos en la recomendación técnica ETSI TR 101 154[38].
2. El sistema MPEG-2 no debe ser utilizado para la codificación de señales de TVAD, ya que está al final de su vida útil y por tanto haría los servicios de TVAD costosos, poco atractivos y con muchas dificultades para encontrar espectro utilizable.
3. El sistema de codificación MPEG-4 AVC/H.264 proporciona una mejora de eficacia de 2:1 con respecto a MPEG-2 (horizonte 2008), y tiene una amplia capacidad de evolución futura y de incorporación de nuevos desarrollos.
4. Los servicios de TVAD deben utilizar como codec de video el estándar MPEG-4 AVC/H.264 en sus configuraciones **MP@L4** ó **HP@L4**.
5. Estudios realizados en los últimos años [25][26], han determinado que la tasa binaria óptima de ambas configuraciones (MP@L4 y HP@L4) tiene una fuerte dependencia con el formato utilizado (Entrelazado-Progresivo), el tamaño y tecnología del display, así como de la naturaleza del contenido. Es por ello que desde la EBU se recomienda en todos sus informes la utilización en la actualidad del formato progresivo 720p frente al formato 1080i25, y del formato 1080p50 como futuro formato de nueva generación de TVAD.
6. La EBU, en su recomendación TECH 3312 [37] de Febrero del 2006, propone para contenidos no críticos la utilización de tasas binarias con MPEG-4 AVC/H.264 entre **8-10 Mbps**, y de hasta **16Mbps** para contenidos de extrema complejidad. Estima que la mejora en la eficiencia de codificación, unida a la utilización de multiplexión estadística, permita en el futuro la difusión de todo tipo de servicios en el rango de **8-10Mbps**.
7. El sistema de codificación VC-1 podría ser una alternativa a MPEG-4 AVC/H.264, pero su eficacia de codificación es apreciablemente menor. Además, su no utilización hasta la fecha para servicios de radiodifusión en Europa, obliga a no recomendar su uso para TVAD en España.
8. Se recomienda encarecidamente la utilización de multiplexión estadística entre programas de TVAD, con el objetivo de aumentar la eficiencia espectral.
9. Para la difusión de servicios con audio multicanal se recomienda la utilización de cualquiera de los dos codec descritos AC3/E-AC3 o HE-AAC v2. Siendo preferible la utilización del segundo en los supuestos en los que el ancho de banda disponible sean inferiores a 192Kbps. Para tasas binarias superiores a los 400Kbps, AC3 y E-AC3 ofrecen una calidad excelente en configuraciones 5.1 con independencia de la naturaleza de los contenidos.

ANEXO 1: BIBLIOGRAFÍA Y REFERENCIAS

- [1] BT.601-5. Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios
- [2] BT.709-5. Valores de los parámetros de la norma de TVAD* para la producción y el intercambio internacional de programas
- [3] UIT-T, H.264. Advanced video coding for generic audiovisual services. ISO/IEC 14496-10:2005, Information technology – Coding of audiovisual-objects – Part 10: Advanced Video Coding
- [4] BT.1120-5. Digital interfaces for HDTV studio signals
- [5] SMPTE 292M-2008. 1.5 Gb/s Signal / Data Serial Interface
- [6] SMPTE 424M-2006. Television – 3 Gb/s Signal/Data Serial Interface
- [7] SMPTE 377M-2004. Television – Material Exchange Format (MXF) – File Format Specification
- [8] SMPTE 384M-2005 Television – Material Exchange Format (MXF) – Mapping of Uncompressed Pictures into the Generic Container
- [9] UIT-R BT.1367* Serial digital fibre transmission system for signals Conforming to Recommendations UIT-R BT.656, UIT-R BT.799 and UIT-R BT.1120
- [10] SMPTE 260M-1999 (Archived 2004) Television – 1125/60 High- Definition Production System – Digital Representation and Bit-Parallel Interface
- [11] SMPTE 240M-1999 (Archived 2004) Television – 1125-Line High- Definition Production Systems – Signal Parameters
- [12] SMPTE 295M-1997 Television – 1920 x 1080 50-Hz – Scanning and Interface
- [13] SMPTE 274M-2008 Television – 1920 x 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates
- [14] SMPTE 296M-2001 Television – 1280 x 720 Progressive Image Sample Structure
- [15] RP 211-2000 (Archived 2006) Implementation of 24P, 25P and 30P Segmented Frames for 1920 x 1080 Production Format
- [16] SMPTE 292-2008 1.5 Gb/s Signal / Data Serial Interface
- [17] SMPTE 372M-2002 Television – Dual Link 292M Interface for 1920 x 1080 Picture Raster
- [18] SMPTE 297-2006 Television – Serial Digital Fiber Transmission System for SMPTE 259M, SMPTE 344M, SMPTE 292 and SMPTE 424M Signals
- [19] SMPTE 425M-2006 Television – 3 Gb/s Signal/Data Serial Interface – Source Image Format Mapping
- [20] SMPTE 435-1-2007 10 Gb/s Serial Signal/Data Interface – Part 1: Basic Stream Distribution
- [21] EBU TECH 3299. *"High Definition image formats for Television Production"*
- [22] EBU Statement D97_revised-2005. HDTV PROGRAMME PRODUCTION: THE NEED FOR FORMAT-AGILE TELEVISION PRODUCTION EQUIPMENT

* Radiocommunication Study Group 6 made editorial amendments to this Recommendation in 2003 in accordance with Resolution UIT-R 44.

- [23] EBU R115-2005. FUTURE HIGH DEFINITION TELEVISION SYSTEMS: The need to develop television production equipment for a progressively scanned image format of 1920 horizontal by 1080 vertical resolution at 50 and 60 Hz frame rates
- [24] EBU Technical Recommendation R112 – 2004 EBU statement on HDTV standards
- [25] "HDTV EBU format comparisons at IBC-2006"
- [26] "Studies on the Bit Rate Requirements for a HDTV Format With 1920 1080 pixel Resolution, Progressive Scanning at 50 Hz Frame Rate Targeting Large Flat Panel Displays "
- [27] ISO/IEC 13818-2 Video - Códice para señales de vídeo entrelazado y no entrelazado
- [28] RP 2008-2008 Material Exchange Format – Mapping AVC Streams into the MXF Generic Container
- [29] ISO/IEC 15441-1. JPEG 2000, Core Coding System (Part 1)
- [30] UIT-R BS.1548-1. User requirements for audio coding systems for digital broadcasting
- [31] UIT-R BT.1737. Use of the UIT-T Recommendation H.264 (MPEG-4/AVC) video source-coding method to transport HDTV programme material
- [32] SMPTE 421M-2006 Television – VC-1 Compressed Video Bitstream Format and Decoding Process
- [33] ETSI TS 102 366 v1.1.1, "Digital Audio Compression (AC-3, Enhanced AC-3) Standard", 2005
- [34] ISO/IEC 14496-3:2005 /Amd.2 Information technology -- Coding of audio-visual objects -- Part 3: Audio
- [35] ETSI TS 102 005 Digital Video Broadcasting (DVB); Specification for the use of video and audio coding in DVB services delivered directly over IP
- [36] "Report of Joint EBU-EICTA Workshop on HDTV audio", Geneva, 29 June 2007
- [37] EBU-TECH 3312 "Digital Terrestrial HDTV Broadcasting in Europe"
- [38] ETSI TR 101 154 v1.8.1. Digital Video Broadcasting (DVB); "Implementation guidelines for the use of MPEG-2 Systems, Video and Audio in satellite, cable and terrestrial broadcasting applications"

ANEXO 2: GLOSARIO

2K	<i>resolución especial de 1080 líneas por 2048 píxeles/línea</i>
4K	<i>resolución especial de 2160 líneas por 4096 píxeles/línea</i>
3GPP	3 rd Generation Partnership Project
ATSC	Advanced Television Systems Committee
AVC	Advanced video Coding
CIF	Common Image Format
CWDM	Coarse Wavelength Division Multiplexing
DCI	Digital Cinema Initiative
DCT	Discrete Cosine Transform
DTS	Digital Theatre System
DVB	Digital Video Broadcasting
EBU	Eurovision Broadcasting Union
EICTA	European Industry Association for Information Systems, Communication Technologies and Consumer Electronics
ETSI	European Telecommunication Standards Institute
GOP	Group of Pictures
ISDB	Integrated Services Digital Broadcasting System
ISO	International Organization for Standardization
MXF	<i>Material Exchange Format</i>
S-DMB	Satellite Digital Multimedia Broadcasting
SMPTE	<i>Society of Motion Picture and Television Engineers</i>
SVC	Scalable Video Coding
TVAD	Televisión de Alta Definición
UIT	Unión Internacional de Telecomunicaciones