

RECOMENDACIÓN UIT-R BT.1687-1

Reducción de la velocidad binaria de vídeo para la distribución* en tiempo real de aplicaciones de imágenes digitales en pantalla grande destinadas a la exhibición en grandes salas

(Cuestión UIT-R 15/6)

(2004-2006)

Cometido

La presente Recomendación trata sobre la reducción de la velocidad binaria de vídeo para la distribución en tiempo real de aplicaciones de imágenes digitales en pantalla grande (LSDI, *large screen digital imagery*) destinadas a la exhibición en grandes salas e incluye ejemplos de velocidades binarias de señales comprimidas utilizando MPEG-2 y MPEG-4/AVC empleadas en aplicaciones de LSDI. En la Recomendación UIT-R BT.1680 se definen los sistemas de imágenes digitales LSDI en banda base.

La Asamblea de Radiocomunicaciones de la UIT,

considerando

- a) que conviene que la distribución y la exhibición de programas de imágenes digitales en pantalla grande (LSDI) mantenga el contenido creativo y la calidad de la imagen del original de distribución¹ en la medida en que lo permitan las condiciones previstas de exhibición del programa;
- b) que la Recomendación UIT-R BT.1680 recomienda sistemas de imágenes digitales que cumplen los requisitos de las aplicaciones LSDI para exhibiciones en grandes salas y que deben utilizarse en la distribución de dichas aplicaciones;
- c) que en la distribución de las aplicaciones LSDI es necesario aplicar una reducción de la velocidad binaria a las señales de programa, a fin de conservar dicha velocidad binaria y de esta manera reducir el tiempo y el costo de la transmisión, sin que ello afecte adversamente a la calidad del programa que percibe el espectador;
- d) que el sistema de visión humano tiene una resolución muy superior para los detalles de luminancia que para los detalles de color;
- e) que el Grupo ISO/CEI JTC 1/SC 29/WG 11 ha efectuado y documentado amplios estudios sobre la velocidad binaria requerida para dar una compresión de vídeo transparente² con diversos algoritmos de compresión y varios sistemas de imágenes³;

* La UIT define la distribución de la siguiente manera: «Envío de programas de televisión cuando no se prevé ningún otro procesamiento post-producción».

¹ La expresión «original de distribución» se utiliza aquí para indicar la copia principal del programa que se obtiene del original terminado tras haberlo adaptado (en contenido y en calidad) al medio de distribución.

² La UIT define la «reducción transparente de la velocidad binaria» como «un proceso de reducción de la velocidad binaria que no afecta a la calidad subjetiva de la secuencia de sonido o de imagen».

³ Véase por ejemplo el Documento ISO/IEC JTC 1/SC 29/WG 11 – MPEG2003/N6231, diciembre de 2003, Waikoloa – Report of The Formal Verification Tests on AVC (ISO/IEC 14496-10 – ITU-T Rec. H.264).

- f) que esos ensayos han confirmado que la codificación MPEG-2 en MP@HL con un codificador⁴ optimizado a 20 Mbit/s es virtualmente transparente a la calidad de la fuente para los sistemas de imagen que se indican en la Recomendación UIT-R BT.1680, con una velocidad de imágenes de 25 ó 30 Hz y exploración progresiva o entrelazada;
- g) que la compresión de vídeo MPEG-2 está muy extendida hoy en día y cumple con la política en materia de patentes de la UIT;
- h) que recientemente se creó un nuevo sistema de compresión de vídeo derivado de MPEG-2 denominado MPEG-4/AVC (véase la Recomendación UIT-T H.264) y que ofrece, aun utilizando los códecs básicos, una eficacia de compresión que es cerca del doble que la de MPEG-2;
- j) que muy recientemente se elaboraron algunas extensiones del sistema original de codificación de fuente MPEG-4/AVC denominadas extensiones de la gama de fidelidad (FRExt), que ofrecen una eficacia de codificación aún mayor y soportan la codificación del sistema de imágenes de 1080/1920p a 50 y 60 cuadros/s;
- k) que la utilización del sistema de codificación de fuente MPEG-4/AVC puede resultar conveniente para las aplicaciones LSDI, ya que requiere una velocidad binaria neta menor para la distribución de programas de LSDI, a pesar que los códecs de MPEG-4/AVC son más complejos y podría por lo tanto esperar que fueran algo más costosos;
- l) que MPEG-4/AVC permite una gran flexibilidad en la elección de los parámetros de codificación de imagen tales como el muestreo 4:4:4, 10 ó 12 bits/muestra y velocidades de cuadro superiores a 72 Hz;
- m) que varios países están implementando actualmente servicios para la presentación de programas LSDI destinados a la exhibición en grandes salas o que prevén implementarlos en el futuro próximo,

recomienda

- 1 que las primeras implementaciones de las aplicaciones LSDI para la exhibición en grandes salas utilicen una señal de vídeo de banda de base digital que se ajuste a la Recomendación UIT-R BT.1680 en las interfaces de entrada y de salida de la cadena de distribución LSDI;
- 2 que, dada la mayor resolución del sistema visual humano a la luminancia que a los detalles de color, se utilice preferentemente la codificación vídeo 4:2:0 o preferentemente la codificación vídeo 4:2:2 para la distribución de programas LSDI destinada a la exhibición en grandes salas;
- 3 que, como la señal de programa LSDI no pasará normalmente por nuevas etapas de postprocesamiento creativo de la imagen tras su distribución, debe darse prioridad a la utilización de la reducción de la velocidad binaria intertrama para la distribución de programas LSDI y su exhibición en grandes salas, ya que ofrece una mayor eficacia de compresión que la codificación interior a la trama; se reconoce que, cuando se requiere la inserción del programa local (EDL o lista de representación) es preferible la codificación interior a la trama y que el valor mínimo de la velocidad binaria de vídeo puede ser superior;
- 4 que, como la reducción de la velocidad binaria utilizada en el canal de distribución debe ser virtualmente transparente a la calidad de la copia original de distribución en las condiciones de presentación previstas, debe considerarse a corto plazo el método de reducción de la velocidad

⁴ La expresión «codificador optimizado» se utiliza aquí para indicar un codificador que cumple la MPEG-2 implementado con circuitos y que incorpora soluciones avanzadas, optimizado respecto al codificador del «modelo de referencia» MPEG-2 TM5 que se implementa con programación y que fue aceptado por el Grupo ISO/CEI JTC 1/SC 29/WG 11 como referencia de calidad.

binaria intertrama MPEG-2 a MP@HL (HiQ) y a una velocidad binaria neta mínima de vídeo del orden de 20 Mbit/s para la distribución en tiempo real de programas LSDI destinados a la exhibición en grandes salas (la utilización de la codificación 4:2:2 @ HL exigirá una velocidad binaria ligeramente superior);

5 que la utilización del método de reducción de la velocidad binaria MPEG-4/AVC se debe también considerar como una alternativa interesante, en particular si se desea utilizar una velocidad binaria neta mucho menor para la distribución de programas LSDI y es factible utilizar códecs algo más complejos;

NOTA 1 – La Recomendación UIT-T H.264 se puede descargar electrónicamente en la dirección: <http://www.itu.int/md/R03-SG06-C-0211/es>.

6 que debe analizarse, en su debido momento, la posibilidad de utilizar los códecs de compresión incluso más eficaces que puedan aparecer en el futuro, si cumplen la política de patentes de la UIT y son probados exhaustivamente, se encuentran enteramente disponibles y preferentemente funcionan con códecs MPEG-2 y MPEG-4/AVC.

Apéndice 1 (Informativo)

Ejemplos de parámetros y herramientas mínimas para codificar en la fuente algunos miembros de sistemas de imágenes de la familia LSDI descritos en la Recomendación UIT-R BT.709, utilizando MPEG-4/AVC (Recomendación UIT-T H.264)

En este Apéndice se sugieren algunos ejemplos de parámetros y herramientas mínimas del método de codificación de la fuente MPEG-4/AVC (Recomendación UIT-T H.264) que podrían utilizarse para comprimir varios miembros de los sistemas de imágenes de la familia LSDI especificados en la Recomendación UIT-R BT.709-5 (Parte 2). El Apéndice también da indicación de las velocidades binarias necesarias para el transporte virtualmente transparente de dichas señales una vez realizada la codificación de fuente.

Miembro de la familia de la Rec. UIT-R BT.709	Herramientas mínimas y parámetros de MPEG-4/AVC	Velocidad binaria para el transporte virtualmente transparente
1 920 × 1 080 × 24/25p	<p style="text-align: center;">Nivel 4</p> <p>Herramientas de codificación Alto 10</p> Principales herramientas de perfil X Formato cromático 4:2:0 X Profundidad de bit de muestra de 8 bits X Adaptabilidad transformada 8 × 8 contra 4 × 4 X Matrices de cambio de escala de cuantificación X Control QP independiente para C_B y C_R X Formato de vídeo monocromático X Profundidad de bit de muestra de 9 y 10 bits X Formato cromático 4:2:2 Profundidad de bit de muestra de 11 y 12 bits Formato cromático 4:4:4 Transformada de color residual Codificación predictiva sin pérdidas	9-11 Mbit/s – Calidad de contribución 7-10 Mbit/s – Calidad de distribución
1 920 × 1 080 × 60/50i	<p style="text-align: center;">Nivel 4</p> <p>Herramientas de codificación Alto 4:2:2</p> Principales herramientas de perfil X Formato cromático 4:2:0 X Profundidad de bit de muestra de 8 bits X Adaptabilidad transformada 8 × 8 contra 4 × 4 X Matrices de cambio de escala de cuantificación X Control QP independiente para C_B y C_R X Formato de vídeo monocromático X Profundidad de bit de muestra de 9 y 10 bits X Formato cromático 4:2:2 X Profundidad de bit de muestra de 11 y 12 bits Formato cromático 4:4:4 Transformada de color residual Codificación predictiva sin pérdidas	10-15 Mbit/s – Calidad de contribución 8-12 Mbit/s – Calidad de distribución

Miembro de la familia de la Rec. UIT-R BT.709	Herramientas mínimas y parámetros de MPEG-4/AVC	Velocidad binaria para el transporte virtualmente transparente
1 920 × 1 080 × 60/50p	<p style="text-align: center;">Nivel 4.2</p> <p>Herramientas de codificación Alto 4:2:2</p> <p>Principales herramientas de perfil X</p> <p>Formato cromático 4:2:0 X</p> <p>Profundidad de bit de muestra de 8 bits X</p> <p>Adaptabilidad transformada 8 × 8 contra 4 × 4 X</p> <p>Matrices de cambio de escala de cuantificación X</p> <p>Control QP independiente para C_B y C_R X</p> <p>Formato de vídeo monocromático X</p> <p>Profundidad de bit de muestra de 9 y 10 bits X</p> <p>Formato cromático 4:2:2 X</p> <p>Profundidad de bit de muestra de 11 y 12 bits</p> <p>Formato cromático 4:4:4</p> <p>Transformada de color residual</p> <p>Codificación predictiva sin pérdidas</p>	<p>18-20 Mbit/s – Formato de emisión</p> <p>NOTA – La utilización de 10 bits/muestra no debe causar un incremento de velocidad binaria por encima de la necesaria para 8 bits.</p>
1 920 × 1 080 × 24/25p	<p style="text-align: center;">Nivel 4</p> <p>Herramientas de codificación Alto 4:2:2</p> <p>Principales herramientas de perfil X</p> <p>Formato cromático 4:2:0 X</p> <p>Profundidad de bit de muestra de 8 bits X</p> <p>Adaptabilidad transformada 8 × 8 contra 4 × 4 X</p> <p>Matrices de cambio de escala de cuantificación X</p> <p>Control QP independiente para C_B y C_R X</p> <p>Formato de vídeo monocromático X</p> <p>Profundidad de bit de muestra de 9 y 10 bits X</p> <p>Formato cromático 4:2:2 X</p> <p>Profundidad de bit de muestra de 11 y 12 bits</p> <p>Formato cromático 4:4:4</p> <p>Transformada de color residual</p> <p>Codificación predictiva sin pérdidas</p>	<p>8-10 Mbit/s – Emisión</p> <p>Modo de «películas» de alta calidad</p>

Adjunto 1

Información sobre el método de codificación de la fuente MPEG-4/AVC [Sullivan y otros, 2004]

La Recomendación UIT-T H.264/MPEG-4 (Parte 10), Codificación de vídeo avanzada para los servicios audiovisuales genéricos (conocida comúnmente como H.264/AVC), es la más reciente de las normas internacionales para codificación de vídeo. Es actualmente la norma más poderosa y actualizada y fue elaborada por el equipo mixto de vídeo (JVT) compuesto por expertos del Grupo de Expertos en codificación de vídeo (VCEG) de la UIT y del Grupo de Expertos en imágenes en movimiento (MPEG) de la ISO/CEI.

De la misma forma como sucedió con las normas anteriores, su diseño ofrece el equilibrio más actualizado entre eficiencia de codificación, complejidad de implementación y coste, teniendo en cuenta los últimos avances de la tecnología de diseño de integración en muy gran escala (VLSI), como por ejemplo las unidades centrales de procesamiento, el procesamiento de señales digitales, circuitos integrados de uso específico y disposición de puertos programables de campo entre otros.

En el proceso se creó una norma que mejoraba la eficacia de codificación en un factor de dos como mínimo (en promedio) con respecto a MPEG-2 al tiempo que el coste se mantenía a niveles aceptables.

En julio de 2004, se añadió una nueva modificación a esta norma, denominada – Extensiones de la gama de fidelidad (FRExt, Enmienda 1), que ha demostrado una eficacia de codificación aún mayor en comparación con MPEG-2, que podría llegar a ser de 3:1 en algunas aplicaciones importantes.

A pesar de su amplia gama de aplicación, la norma original H.264/AVC (completada en mayo de 2003) se centraba principalmente en vídeo con «calidad de los programas de entretenimiento», con 8 bits/muestra y muestreo cromático 4:2:0. Debido a las limitaciones de tiempo, en la norma no se incluyó soporte para los entornos profesionales más exigentes y el diseño no se centró en las resoluciones más altas de vídeo. Para aplicaciones como la contribución de programas y la edición y postprocesamiento de estudio, puede ser necesario:

- Utilizar una precisión de la muestra del vídeo fuente de más de 8 bits por muestra.
- Utilizar, para la representación de color, una resolución mayor que la que comúnmente se utiliza en las aplicaciones de usuario (es decir utilizar muestreo de 4:2:2 o de 4:4:4 en vez del formato de muestreo cromático 4:2:0).
- Llevar a cabo funciones de edición como la combinación alfa (se trata de un proceso mediante el cual se combinan varias escenas de vídeo, y que es muy utilizado en los informes del tiempo en los que la imagen del periodista se superpone sobre la de un mapa o la de un gráfico de radar meteorológico).
- Utilizar velocidades binarias muy altas.
- Utilizar resolución muy alta.
- Alcanzar fidelidad muy alta e incluso representar algunas partes del vídeo sin pérdidas.
- Evitar errores de redondeo en la transformación color-espacio.
- Utilizar la representación de colores rojo, verde, blanco (RGB, *red, green, bleu*).

El proyecto FRExt dio como resultado un conjunto de cuatro nuevos perfiles conocidos colectivamente como perfiles altos:

- El perfil alto (HP), que soporta vídeo de 8 bits con muestreo 4:2:0, se destina a usuarios finales con alta exigencia y a otras aplicaciones que utilizan una alta resolución de vídeo sin necesidad de formatos cromáticos extendidos ni de exactitud de muestra extendida.

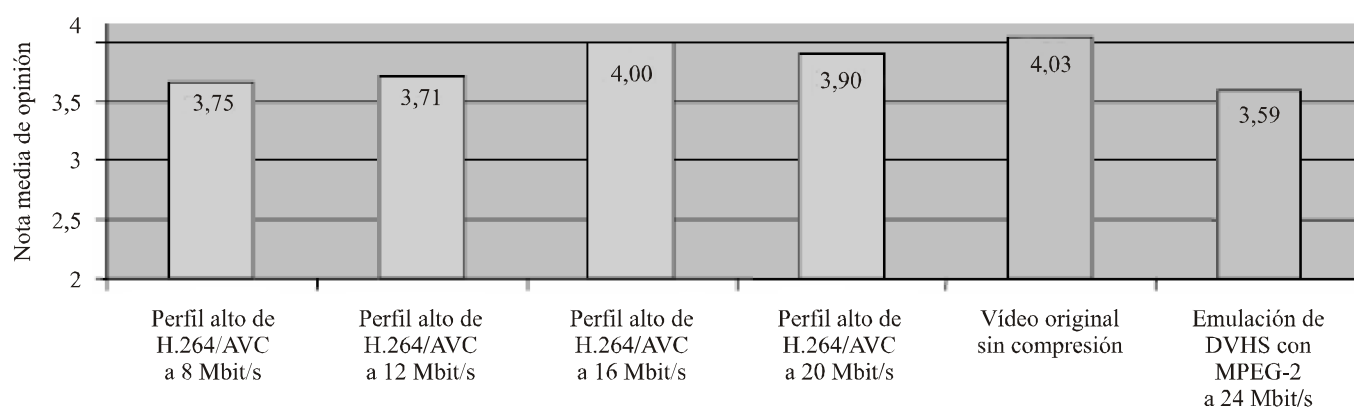
- El perfil alto 10 (Hi10P), que soporta vídeo 4:2:0 con hasta 10 bits de precisión de representación por muestra.
- El perfil alto 4:2:2 (H422P), que soporta muestreo cromático 4:2:2 y hasta 10 bits por muestra.
- El perfil alto 4:4:4 (H444P), que soporta muestreo cromático 4:4:4, hasta 12 bits por muestra y soporta adicionalmente codificación eficiente sin pérdidas y una transformada de color residual entera al codificar vídeo RGB, al tiempo que se evitan errores en la transformación color-espacio.

Como FRExt es bastante nueva y algunos de sus beneficios no son objetivos sino que se dan desde el punto de vista de la percepción, es algo más difícil medir sus posibilidades. Un punto de referencia pertinente son los resultados de la evaluación subjetiva de calidad realizada por la Blu-ray Disc Association (BDA). Los resultados se reproducen en la Fig. 1 a partir del informe de pruebas [Wedi y Kashiwagi, 2004]. Esta prueba, que se realizó utilizando una película de 24 cuadros/s con exploración progresiva de $1\,920 \times 1\,080$, arroja los siguientes resultados nominales (que no se deben considerar como comprobados con rigurosidad estadística):

- Con el perfil alto de FRExt se obtuvo una calidad de vídeo nominalmente mejor que la de MPEG-2, utilizando sólo una tercera parte de los bits (8 Mbit/s contra 24 Mbit/s).
- Con el perfil alto de FRExt se obtuvo una calidad de vídeo nominalmente transparente (es decir que es difícil diferenciarla del vídeo original sin compresión) a tan sólo 16 Mbit/s.

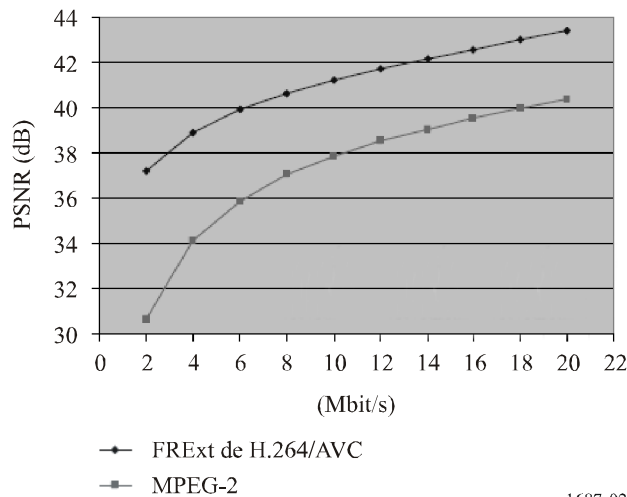
Se sobrepasó significativamente la barra de calidad (3.0), cuyo uso se considera adecuado en esta organización para medios empaquetados de alta definición, utilizando sólo 8 Mbit/s. Además, como en estas pruebas no se optimizó el método de codificación H.264/AVC utilizado, posiblemente se podría reducir significativamente la velocidad binaria a menos de 8 Mbit/s sin bajar de la barra de calidad 3.0, para lograr así una calidad lo suficientemente alta como para decir que se trata de una «alta definición aceptable» para esa aplicación exigente.

FIGURA 1
Comparación de MPEG-2 con H.264



En la Fig. 2 se muestran los resultados de un ejemplo de prueba comparativa objetiva (relación entre señal de potencia y ruido, PSNR) realizada por FastVDO⁵. Estos resultados objetivos confirman el buen comportamiento del perfil alto. (También en este caso, la utilización no óptima de cuadros B hace que el comportamiento de FRExt que se refleja en el gráfico resulte algo conservador.)

FIGURA 2
Comparación de PSNR



1687-02

Referencias Bibliográficas

- SULLIVAN, G.J., TOPIWALA, P. y LUTHRA, A., [2004] The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extension. Presentado en la SPIE Conference on Applications of Digital Image Processing XXVII. Special Session on Advances in the New Emerging Standard: H.264/AVC.
- WEDI T. y KASHIWAGIY., [2004] Subjective quality evaluation of H.264/AVC FRExt for HD movie content. Documento del Equipo mixto de vídeo JVT-L033.

⁵ La empresa FastVDO está especializada en tecnología para comunicación de medios y programas informáticos de infraestructuras. Está ubicada en Columbia, MD (Estados Unidos de América).