

LA COLUMNA DE ANDRES MAYO

HISTORIA, PRESENTE Y FUTURO DE LA CODIFICACION DE AUDIO MULTICANAL

LAS NECESIDADES DE NUESTRA INDUSTRIA HACEN QUE EL AUDIO MULTICANAL CASI SIEMPRE DEBA COMPRIMIRSE PARA DEJAR EL MÁXIMO ESPACIO POSIBLE AL VIDEO. AQUÍ, UN PANORAMA DE SITUACIÓN DE LOS ESQUEMAS MÁS COMUNES DE CODIFICACIÓN.

Los que hemos tenido la experiencia de trabajar en formatos multicanal (típicamente para DVD-Video), sabemos que una de las partes más difíciles de nuestra tarea es la de llevar al disco el resultado de nuestro trabajo conservando al máximo su calidad original, es decir logrando que la compresión no afecte perceptiblemente el producto final. Esto resulta sumamente complejo en algunos casos por la impiadosa relación de compresión que necesitamos aplicar según el espacio disponible.

Muchas veces, al preparar el audio para un DVD se me presenta la incógnita: ¿qué formato final sería el más adecuado y porqué? Existiendo varias posibilidades, lo lógico resulta comenzar por los formatos menos comprimidos (wav para el audio stereo y DTS full para el surround 5.1). Para que estos dos streams de audio puedan coexistir con el video adentro de un DVD de unos 90 minutos de duración precisamos que el disco tenga un espacio disponible (adicional al espacio ocupado por el video, subtítulos, animaciones, menús, etc) de unos 3 GB, lo cual lleva inevitablemente a la necesidad de producir un DVD dual-layer (disco de doble capa, que almacena 8.54 GB en lugar de los 4.7 GB de un DVD single-layer). Esto implica un mayor costo en la realización del stamper y también en la duplicación, por lo cual debe ser previsto en la etapa de preproducción. La mayoría de las veces, el espacio disponible en el DVD no es suficiente (aún para un dual-layer) y termino optando por entregar el audio en AC-3, incluso para el stereo. Es importante agregar que el AC-3 o Dolby Digital (originalmente desarrollado para el cine) hace un muy buen trabajo dadas las difíciles condiciones de codificación impuestas, porque su "bit rate" (es decir la cantidad de bits por segundo máxima de que puede disponer) es de apenas 448 Kbps. 3 veces menor que en el DTS full!!

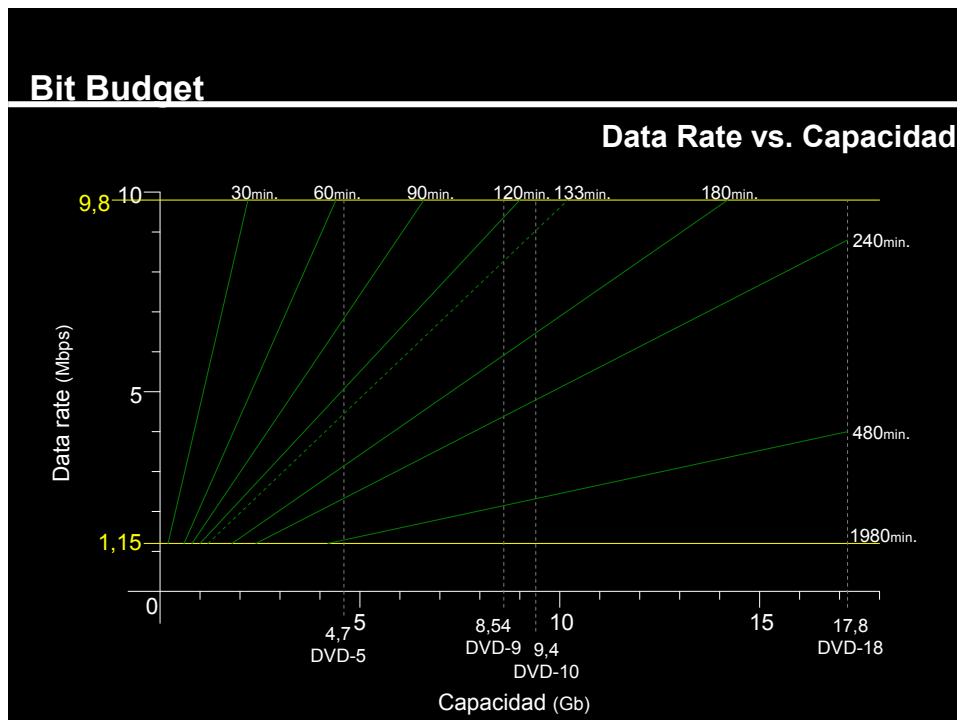
Por norma, el bit rate máximo que puede soportar el lector del DVD es de 9.8 Mbps, que no debe ser superado en ningún momento de la reproducción porque sencillamente el lector se verá excedido en su capacidad. Este bit rate se reparte entre el video (gran consumidor de bits por segundo!!), el audio, los subtítulos y todos los demás elementos simultáneos que deba reproducir el láser. El "consumo" de bit rate varía enormemente según la norma de codificación empleada, desde los módicos 192 Kbps para un audio stereo codificado en AC-3 (con alta compresión) hasta los más de 1536 Kbps de un audio stereo codificado en WAV (sin compresión *). Necesariamente las diferencias tienen que ser perceptibles entre uno y otro.

* Cabe la aclaración de que el audio estará siempre *codificado*, pero no necesariamente siempre *comprimido*.

Lo que en mi caso determina la elección de DTS o AC-3 es el espacio y el bit rate disponibles para el audio: entonces, al codificar la mezcla 5.1 de un DVD musical, si tengo mayor espacio y bit rate utilizo DTS y si estoy comprometido en alguno de estos dos parámetros utilizo AC-3. En la FIGURA 1 se pueden ver las distintas capacidades de los DVD de acuerdo con su tamaño. Para ubicarnos en el gráfico, un punto de referencia es el cruce entre la línea vertical punteada de 4.7 GB correspondiente al DVD-5 (o single-layer) y la línea punteada diagonal correspondiente a los 133 minutos de programa (máximo teórico recomendado para un DVD-5). Según esta referencia, el bit rate promedio para todo el DVD es inferior a los 5 Mbps. Este bit rate es apenas suficiente para codificar un material de video que no presenta demasiadas exigencias y cuyo audio está codificado en AC-3. Si queremos agregar una opción de audio en DTS o WAV stereo, este bit rate resulta claramente insuficiente y tendremos dos opciones:

reducir la cantidad de minutos de video en el DVD o bien pasar a la siguiente versión: el DVD-9 (dual-layer), con una capacidad teórica de 8.54 GB. Este cambio hará correr la línea punteada vertical hacia la derecha y el nuevo punto de cruce con los 133 minutos de nuestro programa de video arroja un bit rate por encima de los 8 GB.

FIGURA 1: BIT RATE VS. CAPACIDAD EN LOS DISTINTOS TAMAÑOS DE DVD



Veamos un poco de historia sobre estos formatos:

Ambos estándares fueron creados en pos de ofrecer a los consumidores una calidad de audio capaz de satisfacer sus siempre crecientes demandas, que a partir del CD se situaron bien por encima de los 80 dB de rango dinámico y relación señal-ruido y más de 15 KHz de contenido espectral de la señal de audio. Todas las entidades creadoras de estándares, incluyendo a AES, SMPTE, EBU, ITU, MPEG, etc. trabajaron a fines de la década del '80 y comienzos de los '90 para finalmente converger en un único standard denominado comunmente 5.1 (o también 3/2/.1, en referencia a la ubicación de las bocinas: 3 adelante, 2 atrás y 0.1 para la reproducción de las bajas frecuencias). Desde entonces, muchos adelantos notables en este campo impulsaron diferentes esquemas de sonido envolvente: el ya muy utilizado 7.1, el 10.2 (con bocinas en altura y doble sub-woofer), el 22.2, etc. Obviamente cada uno de estos avances representó un gran desafío desde el punto de vista de la codificación del audio, porque seguían existiendo las limitaciones de espacio y bit rate antes descritas.

Por estas mismas razones no resultaba práctico utilizar para el audio multicanal la codificación conocida como Linear PCM (que es la utilizada por el CD de audio convencional), aún a pesar de ser ampliamente probado y sumamente económico de implementar. Una simple cuenta nos demuestra que el bit rate necesario para 5.1 canales en PCM (utilizando apenas 44.1 KHz de frecuencia de muestreo y 16 bits de profundidad) es de casi 3.6 Mbps, muy por encima de los máximos permitidos por estándares actuales como la televisión digital, Internet, etc. Por lo tanto la codificación multicanal en Linear PCM quedó relegada a usos no masivos como el DVD-Audio. Más aún, la tecnología PCM había adoptado una distribución uniforme de bits a lo largo de todo el espectro sonoro, en tanto que numerosas investigaciones realizadas en los últimos 20 años demuestran que una distribución de bits modelada de acuerdo con las características del oído humano mejoran significativamente la calidad de audio percibida, ya que algunos

componentes espectrales son mucho más audibles que otros, y contribuyen a disimular los ruidos de cuantización en frecuencias medias producidos por los convertidores A/D. Sobre este tema, puede leerse el excelente trabajo presentado por Louis Fielder en la AES en 1987.

Todos los esquemas de codificación modernos tienen la capacidad de eliminar las redundancias en la señal y redistribuir el "paquete" de bits de la manera más apropiada en el dominio de la frecuencia, utilizando mayor cantidad de bits en donde se hace más audible el efecto de la compresión, a expensas de quitar bits en donde no son tan necesarios. Incluso se aplican técnicas complejas de reducción de ancho de banda como el llamado Channel Coupling, que consiste en utilizar una señal espectral combinada en lugar de transmitir la información espectral de cada canal por separado. En un par stereo, esta señal combinada puede obtenerse mediante una aproximación tipo M/S, en la cual el contenido espectral de cada canal es sumado y restado entre sí y dependiendo del grado de correlación que exista entre ambos se utilizará la señal original o la suma-diferencia de ellos.

La diferencia fundamental entre DTS y AC-3 estuvo dada desde el origen por el hecho de que el AC-3 fue concebido como parte integrante del film de 35 mm utilizado en el cine, por lo tanto su bit rate máximo está determinado por la cantidad de bits que pueden almacenarse en el pequeño espacio disponible entre las inserciones o *sprockets*. En cambio, el DTS se lee desde un CD-ROM externo que tiene una capacidad y velocidad de transferencia muy superior (ver FIGURA 2). Este esquema permite que el DTS pueda almacenar 5.1 canales de audio con una relación de compresión de sólo 4:1, muy baja en comparación con la del AC-3 en el que la compresión es superior al 10:1. Otro estándar muy utilizado es el desarrollado por ISO/IEC, conocido como MPEG-2 AAC (Advanced Audio Coding) que fue la base del MPEG-4. Esta es la tabla comparativa de parámetros aceptables por cada estándar:

AC-3

Cantidad de canales de audio: 1 a 5.1
Frecuencias de muestreo aceptadas: 32, 44.1, 48 kHz
Profundidades de bits aceptadas: 16, 24
Ancho de banda: 32 a 640 Kbps
Tamaño del frame: 1536 samples.

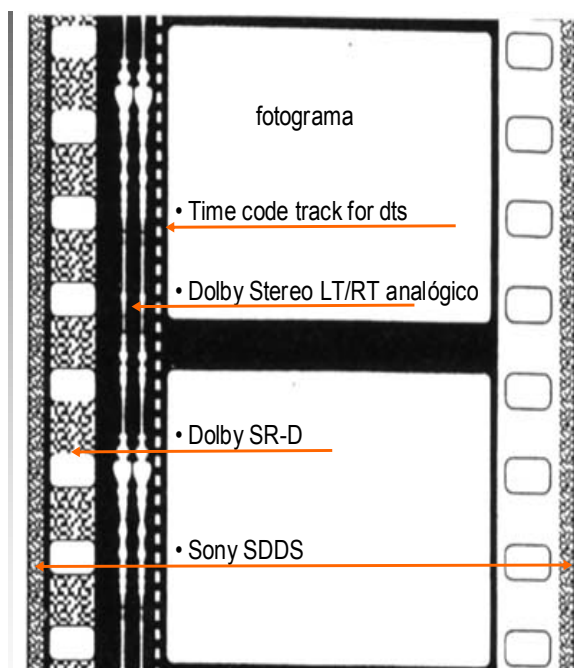
DTS

Cantidad de canales de audio: 1 a 10.1
Frecuencias de muestreo aceptadas: 8 a 192 KHz
Profundidades de bits aceptadas: 16, 24
Ancho de banda: 32 a 6144 Kbps
Tamaño del frame: 512 samples.

AAC

Cantidad de canales de audio: 1 a 48
Frecuencias de muestreo aceptadas: 8 a 96 KHz
Profundidades de bits aceptadas: 16, 24
Ancho de banda: hasta 576 Kbps
Tamaño del frame: 1024 samples.

FIGURA 2: ALOJAMIENTO DE LOS DISTINTOS FORMATOS DE AUDIO EN EL FILM DE 35 mm



Para los entendidos, interesa agregar que el mapeo del tiempo en función de la frecuencia en el AC-3 se implementa mediante un banco de filtros variable en el tiempo, que utiliza una transformada de coseno modificada (MDCT) de 128 puntos para los transientes y de 256 puntos para las señales estables. Muy similar al esquema empleado por el algoritmo AAC, sólo que este último utiliza una MDCT de 1024 puntos para las señales estables. En cambio el esquema DTS divide el espectro en 32 bandas de frecuencia entre 0 y 24 KHz, con 10 sub-bandas adicionales para cubrir el rango hasta los 96 KHz. Aquí el mapeo de tiempo en función de la frecuencia se realiza mediante un filtro tipo PQMF (Polyphase Quadrature Mirror Filter). Para la mayoría de las aplicaciones, DTS aplica relaciones de compresión muy bajas, casi en el rango de la operación conocida como "lossless" (sin pérdidas). En resumen, AC-3 obtiene buenos resultados para codificación de 5.1 canales en el rango de los 450 Kbps y DTS tiene buen resultado a partir de 1 Mbps. En cambio, AAC ha demostrado ser capaz de proveer buenos resultados a bit rates extremadamente bajos, como 64 Kbps, lo cual lo ha colocado en un sitio preferencial a la hora de elegir un esquema de codificación para Internet.

Conclusiones

Es cierto que la capacidad operativa de los procesadores se incrementa a pasos agigantados, y también el ancho de banda disponible a través de Internet o de los nuevos formatos de distribución como el HD-DVD o Blu-Ray. Esto nos podría llevar a pensar en la distribución de audio sin compresión para algunos formatos, pero también es cierto que los algoritmos de compresión de datos seguirán siendo muy importantes para otros casos, especialmente para broadcasting digital, considerando los cada vez más altos requerimientos de cantidad de canales y resolución de audio existentes. Pero lo verdaderamente fundamental es que sigan existiendo artistas, productores e ingenieros que se animen a producir para estos formatos, en un rango que va desde el realismo puro de la recreación de un concierto hasta la invención de dimensiones totalmente nuevas para la composición de una pieza. Las herramientas ya están disponibles, sólo falta jugar, crear y experimentar con ellas.

Este artículo puede descargarse en formato pdf desde www.andresmayo.com/data

Andrés Mayo es ingeniero de Mastering y realizador de DVD musicales.

Miembro fundador de Team 5.1, es reconocido en Argentina por sus trabajos de masterización en stereo y surround Es Vicepresidente de A.E.S. Región América Latina. Contacto: aam@aes.org