

PRECISIONES SOBRE EL FUNCIONAMIENTO DEL DOLBY DIGITAL ®

CUANDO TERMINAMOS DE MEZCLAR Y MASTERIZAR EL AUDIO 5.1 DE UN DVD, COMIENZA LA TAREA DE COMPRIMIRLO EN ALGUNO DE LOS DOS FORMATOS ACEPTADOS: DOLBY DIGITAL (MANDATORIO) O DTS (OPCIONAL). A CONTINUACION, UN RESUMEN DE LA INFORMACION A TENER EN CUENTA PARA UTILIZAR DOLBY DIGITAL

No mucho se conoce acerca de la importancia de la correcta codificación del audio 5.1 para un DVD. A partir de una gran cantidad de situaciones en las que la falta de espacio físico en el disco era uno de los mayores problemas a resolver, se me fue haciendo cada vez más evidente que la compresión de datos en los canales de audio no podía ser dejada librada al azar sino que tenía que ser encarada profesionalmente como una tarea de mucha importancia en el rendimiento final del sonido de un DVD. Algo de esto fue visto en la nota publicada en Abril de 2007 en esta misma revista, pero me gustaría dar ahora algunas precisiones técnicas que pueden ser de utilidad. A pesar de que el DVD tiene desde hace tiempo un hermano más joven y poderoso llamado Blu-Ray y aún considerando las grandes chances de que todo formato físico termine en breve por desaparecer en manos de la virtualidad, en América Latina muchísimos artistas editan sus nuevos álbumes en formato DVD y creo que así será por algunos años todavía.

Comencemos por hacer un breve repaso de la razón de existencia del Dolby Digital: mucho antes de la era digital, el sistema Dolby ya era una parte esencial de la industria del entretenimiento audiovisual porque su sistema de codificación había sido específicamente diseñado para caber en la cinta de 35mm de uso común en el cine. Este desarrollo permitió a Dolby convertirse en el formato estándar de audio en todas las películas, situación que se mantiene hasta el día de hoy. Con la aparición del DVD se buscó la forma de adaptar el formato a los nuevos requerimientos, dando lugar al desarrollo de lo que hoy conocemos como Dolby Digital, originalmente denominado AC3 o Algoritmo de Codificación de tercera generación. Competiendo con formatos mucho más nuevos como el DTS (cuya relación de compresión es notablemente menor y por lo tanto está generalmente considerado como la mejor opción en cuanto a calidad sonora para el audio 5.1 de un DVD), Dolby ha logrado sin embargo ser aceptado universalmente como formato mandatorio (es decir obligatorio) en la etapa de *authoring* del DVD. Esto significa que un DVD con audio en 5.1 puede no ser compatible con DTS, pero sí debe ser compatible con los decodificadores de Dolby. De allí la enorme importancia de conocer a fondo su funcionamiento.

Tomando como punto de partida a la cinta de 35mm, el ancho de banda queda prefijado a partir de la cantidad máxima de bits que pueden alojarse en el pequeño espacio disponible al costado de cada cuadro en el filmico. Los bits deben ser leídos mediante un lector óptico y por lo tanto su tamaño no puede reducirse demasiado porque dejarían de ser legibles, así que no pueden alojarse más bits en ese mismo espacio. Multiplicando esa matriz de bits por la cantidad de cuadros por segundo de la imagen en el cine (24 fps) tenemos 448 Kbps de ancho de banda para el Dolby, lo cual implica una reducción de más de 10 veces con respecto al audio original. Esta es una de las razones poderosas para prestar mucha atención a la compresión de audio, porque efectivamente el espacio ocupado por el mismo se está reduciendo en más del 90%! Lo que se dice una solución muy costo-eficiente, que deja sin embargo sus secuelas en el audio, ya que en un alto porcentaje de los casos la versión comprimida es distinguible de la versión original. Ahora bien, para que esta compresión no tenga efectos negativos demasiado perceptibles, el modelaje matemático desarrollado por Dolby asigna mayor cantidad de bits en la codificación a las señales más audibles y menos bits a las menos detectables: es lo que se denomina una compresión de "bit rate variable".

Para asignar el "bit rate" necesario en cada instante, es necesario hacer un minucioso análisis del contenido del material sonoro. El algoritmo implementado por Dolby Digital (copiado luego por muchos desarrollos similares) divide el espectro disponible (aproximadamente de 20 a 20.000 Hz) en 32 "bandas de frecuencia" bastante angostas, utilizando modelos matemáticos

que buscan copiar la respuesta en frecuencia del oído humano. Profundizando los conocimientos sobre la llamada codificación perceptual, se descubrieron fenómenos que permitieron reducir notablemente la cantidad de información a transmitir en cada instante. Uno de ellos es el que se conoce como enmascaramiento en frecuencia (*frequency masking*), que toma en cuenta el hecho de que (dentro de un mismo rango de frecuencias) el oído humano es menos sensible a los sonidos más suaves cuando hay presencia de otros sonidos más intensos. Por ejemplo, si en un grupo de rock el baterista toca el *hi hat* muy suavemente y al mismo tiempo el *crash* mucho más fuerte, el *hi hat* pasará totalmente desapercibido para el oído común, porque se trata de sonidos que ocupan el mismo rango de frecuencias. Esto no sería cierto si escucháramos simultáneamente el *hi hat* y el bajo, ya que su rango de frecuencias está tan alejado que podemos distinguir ambos sonidos con independencia de sus intensidades, porque no se produce el efecto de enmascaramiento en frecuencia. Otros fenómenos estudiados, como el enmascaramiento temporal y el umbral adaptativo del oído, fueron ampliamente aprovechados para lograr que en cada banda de frecuencia analizada la información resultante sea la que el oído tiene mayores probabilidades de detectar. Precisamente, la calidad de uno u otro formato de compresión radica en el depuramiento de las operaciones matemáticas necesarias para lograr el máximo rendimiento de compresión en cada banda de frecuencia sin alterar perceptiblemente la información audible. Sacando ventaja de estas características de la audición humana, el audio puede ser codificado de manera mucho más eficiente y las secuelas mencionadas se reducen notablemente.

Es muy importante hacer notar que todas estas consideraciones son especialmente ciertas para el espectador promedio, es decir para el oído no entrenado, que constituye el 95% de la gente que consume música. A esta altura, cabría una pregunta: ¿porqué se hace necesario reducir la cantidad de información en los canales de audio? Simplemente porque de otra forma no cabría: en un DVD promedio, más del 80% del espacio disponible está ocupado por el video y el resto se asigna de manera proporcional entre el audio, los subtítulos, menús, etc. Si no comprimiéramos la cantidad de información que ocupa cada canal, no habría espacio físico en el disco para alojar 6 canales de audio discreto y además la velocidad de transferencia de datos requerida por el audio PCM (sumada a la del video) sería imposible de alcanzar por el limitado lector óptico del reproductor de DVD. En el caso del DVD particularmente, la complejidad adicional está dada por el hecho de que hay que comprimir 5.1 canales de audio en menos espacio del que ocuparía un canal PCM mono. Este problema de la falta de espacio y de ancho de banda, afortunadamente ya resuelto en el Blu-Ray, representa un dolor de cabeza para quienes trabajamos en la producción del audio surround y vemos que nuestro producto final sufre una reducción de datos cuya relación de compresión es similar a la de un MP3.

Además existe otro problema importante: ¿cómo determinar si el espectador va a reproducir el audio en forma idónea o no? Para zanjar esta incertidumbre se puso en práctica una idea muy innovadora de Dolby, consistente en crear un sistema de codificación capaz de entregar parámetros de control (metadata) junto con el audio comprimido. Estos parámetros permiten por ejemplo determinar el nivel relativo de reproducción de sonido (conocido como dialnorm o Dialog Normalization), el ajuste de la Compresión de Rango Dinámico (conocida como DRC) y la configuración del *downmix* stereo, es decir cuando se traslada la mezcla a dos canales. Todas estas opciones son deseables cuando se conocen sus efectos pero al mismo tiempo pueden ser peligrosas en manos de quien no domina el tema, provocando una decodificación incorrecta del lado del usuario final y en última instancia una degradación en la calidad de escucha. Hay que tener en cuenta que los espectadores pueden escuchar el audio de un DVD en una gran variedad de formatos, incluyendo los siguientes:

- 5.1 canales con rango dinámico completo
- 5.1 canales con rango dinámico reducido (conocido como Late Night Listening, para quienes no desean molestar a los vecinos con el sonido de explosiones y gritos)
- Downmix estéreo compatible con Dolby Surround
- Downmix estéreo normal
- Downmix mono

Dolby Digital atiende estas necesidades proveyendo múltiples configuraciones de audio, que van desde un canal único (configuración mono) hasta 5.1 canales discretos, con una reducción de cantidad de información entre 10:1 y 12:1. Los 5 canales discretos principales tienen un rango de frecuencias entre 3 Hz (filtrado con un pasaaltos para evitar la superposición de un nivel indeseado de corriente continua) y 20,7 KHz, mientras que el canal discreto destinado al LFE tiene un rango entre 3 Hz y 120 Hz.

Una opción interesante que ofrece el Dolby Digital es la de optimizar el nivel de los canales traseros (Ls y Rs) y Central para ajustar a la preferencia del ingeniero en el downmix. Como la mayoría de los reproductores de DVD tienen ya prefijados de fábrica los parámetros de decodificación referentes al dialnorm, al DRC y al downmixing, es extremadamente importante que el ingeniero a cargo de la codificación utilice un decodificador profesional de tiempo real para monitorear las variables a ajustar. Por supuesto los valores que el ingeniero defina para cada una de estas variables tendrán un impacto directo en el comportamiento del decodificador, o dicho de otra manera: la codificación es esencialmente llevada a cabo por el decodificador de cada DVD player en el momento de la escucha. Veamos entonces cuál es el concepto del dialnorm: cuando pasamos de un tipo de contenido a otro, por ejemplo de un programa de deportes a una banda de rock y luego a una orquesta, las diferencias de presión sonora pueden hacerse muy grandes, del orden de los 20 dB. Se hace necesario entonces medir esta diferencia para poder ajustarla de alguna manera. Para estos fines, el volumen de cada programa sonoro se mide utilizando el método LAeq (Longterm Average A-weighted sound pressure), cuyos resultados aproximan con bastante precisión a la percepción del espectador frente a la pantalla y se miden en dBFS LAeq. Por ejemplo, el nivel estándar con el que los Ingenieros han venido mezclando el diálogo de las películas durante años está en el orden de -25 a -31 dBFS LAeq. Así, todos los programas sonoros pueden ajustarse de forma que mantengan relación adecuada con este volumen de diálogo, dando origen al dialnorm como parámetro de control de volumen. Lo interesante de este parámetro es que es ejecutado por el reproductor en base a una decisión que se toma en la etapa de codificación del audio. La fórmula de la atenuación que necesitamos introducir a un determinado programa sonoro para "empatarlo" con el nivel del diálogo es simplemente: $31 + \text{dialnorm}$, donde dialnorm es un número siempre negativo. Dado que Dolby Digital tiene también gran aplicación en la codificación de películas y otro tipo de programas sonoros, hay muchos casos en los que el valor del dialnorm se prefija en el orden de -15 a -20 (típicamente para broadcasting). Pero en el caso de la música, si estamos conformes con la dinámica de nuestras mezclas y no queremos introducir atenuación al momento de la decodificación, lo correcto es darle al dialnorm un valor de -31. Esta es mi recomendación para la codificación de mezclas 5.1 con buena dinámica, donde no queremos que ningún parámetro de control las modifique. Es importante tener en cuenta que otros parámetros opcionalmente activables como el DRC dependen enteramente del ajuste del dialnorm ya que su cálculo es relativo a este valor y no absoluto. Por lo tanto, ajustar adecuadamente el valor del dialnorm para cada programa sonoro es clave en la calibración de todo el sistema de control de rango dinámico.

Otros temas interesantes relacionados son, por ejemplo, el control de graves (también conocido como Bass Management) y la calibración de niveles en los distintos canales. Por falta de espacio en este artículo recomiendo a quien esté interesado que consulte las Encoding Guidelines de Dolby.

Hasta la próxima!

Ing. Andrés Mayo

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

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

Miembro fundador de Team Surround, es reconocido por sus trabajos de masterización stereo y surround. Es Vicepresidente de A.E.S. Región América Latina.

Contacto: andres@andresmayo.com