

SOBRE EL USO DEL ORDENADOR EN LAS ACTIVIDADES DE OCIO Y TIEMPO LIBRE (COMPRESIÓN DE MÚSICA DIGITAL EN MP3)

Francisco José Muñoz Vivas

RESUMEN

Los archivos en formato MP3 son archivos de música, comprimidos mediante el formato **MPEG**. Los archivos de música comunes (WAV) son comprimidos en gran medida y pasados a formato **MPEG**. Dada la magnitud de la compresión se ha vuelto muy común utilizarlos para comprimir temas enteros de música de CD en formato **MP3 y calidad CD**, los cuales "pasan" de ordenador en ordenador y a través de Internet. El éxito de este formato sobre otros formatos (como el nuevo formato de Yamaha – TVQ) estriba en la gran compresión, las opciones de calidad al comprimir y la facilidad de uso de los programas que existen.

Aunque menos popular el uso de este sistema de compresión se está extendiendo profesionalmente, y con gran rapidez, entre la industria audiovisual, tanto en telecomunicaciones como en ocio .

Palabras claves: Digitalización, CDs, Música comprimida, MP3.

1. - Introducción.

Cada vez es más común el uso del ordenador en casa, a la vez que se extiende y universaliza, no solo para tratamiento de textos y actividades relacionadas con lo puramente profesional.

Las actividades de ocio y tiempo libre se van abriendo paso paulatinamente al usuario en general y es en este contexto en el que se desenvuelven con mayor éxito los formatos de compresión.

El formato MP3 (Mpeg Layer III) no es más que un wav (wave) con una calidad muy elevada de compresión. El formato wav es un formato de sonido (de onda) que abarca distintos grados de calidad de sonido en base (bitrate) a la frecuencia de muestreo (en Hz) así como al sonido estéreo o mono. Con un archivo

wav podemos conseguir una calidad de reproducción igual a la obtenida con un disco compacto pero el problema estriba en la gran cantidad de espacio que ocupan estos archivos. Por medio de compresores se consigue reducir el tamaño de estos archivos de una forma exponencial y así, con el formato comprimido MP3, podemos tener en un poco más de tres Megabytes una canción de tres minutos con calidad CD.

A través de Internet se ha venido extendiendo en el ámbito de la informática doméstica el estándar MPEG-1 layer-3, incorrectamente llamado MPEG-3 o MP3, un esquema de codificación de audio general con una asombrosa capacidad de compresión sin pérdida aparente de calidad, superando una ratio de entre 10 y 12 a 1, sin necesidad de un potente ordenador y con la posibilidad de descomprimir en tiempo real sin usar muchos recursos que nos deje "colgado" el procesador. El único problema es que para poder escuchar las canciones tendrá que hacerse a través del ordenador, con un programa reproductor de MP3, aunque ya han salido al mercado reproductores que se acoplan a la cadena de música e incluso están apareciendo los primeros prototipos para el automóvil. A partir de 1997 estalló el boom de la compresión MP3 y hoy son innumerables los servidores que ofrecen música reducida, a un tamaño manejable. Los CDs tradicionales una vez codificados pueden almacenar ahora hasta 12 horas de audio de alta calidad, las páginas WEB ofrecen sonido «auténtico» en lugar del tradicional MIDI, se ofrecen servicios de música a la carta y programas de radio de calidad superior a la tradicional a través de Internet, y todo ello con muchísima rapidez y calidad dependiendo de la conexión. Todo ello es debido al resultado de comprimir los archivos de audio, y los esquemas de codificación perceptuales, que aprovechan determinadas características de la audición del ser humano para eliminar información innecesaria.

El sonido es una onda continua que se propaga a través del aire u otros medios, y que está formada por diferencias de presión, de forma que puede detectarse por la medida del nivel de presión en un punto. Las ondas sonoras poseen las características propias y estudiadas de las ondas en general, reflexión, refracción y difracción. Al tratarse de una onda continua, se requiere un proceso de digitalización para representarla como una serie de números. Actualmente, la mayoría de las operaciones realizadas sobre señales de sonido son digitales, pues tanto el almacenamiento como el procesado y transmisión de la señal en forma digital ofrece ventajas muy significativas sobre los métodos analógicos. La tecnología digital es más avanzada y ofrece menor sensibilidad al ruido en la transmisión y capacidad de incluir códigos de protección frente a errores, así como posibilidad de encriptación. Con los mecanismos de decodificación adecuados, se pueden tratar simultáneamente señales de diferentes tipos transmitidas por un mismo canal. La desventaja principal de la señal digital es que requiere un ancho de banda mucho mayor que el de la señal analógica,

El proceso de digitalización se compone de dos fases: muestreo y cuantización. En el muestreo se divide el eje del tiempo en segmentos discretos: la frecuencia de muestreo será la inversa del tiempo que medie entre una medida y la siguiente. En estos momentos se realiza la cuantización, que, en su forma más sencilla, consiste simplemente en medir el valor de la señal en amplitud y guardarlo.

Por tanto, siendo el rango superior de la audición humana en torno a los 20 Khz, la frecuencia que garantiza un muestreo adecuado para cualquier sonido audible será de unos 40 Khz. Para obtener sonido de alta calidad se utilizan frecuencias de entre 44'1 Khz, en el caso del CD, por ejemplo, hasta 48 Khz, en el caso del DAT. Otros valores típicos son submúltiplos de la primera, 22 y 11 Khz. Según el uso que se vaya a dar a la compresión, las frecuencias adecuadas pueden ser muy inferiores, de tal manera que el proceso de la voz acostumbra a realizarse a una frecuencia de entre 6 y 20 Khz. o incluso menos. En lo referente a la cuantización, es evidente que cuantos más bits se utilicen para la división del eje de la amplitud, más «fina» será la partición y por tanto menor el error al atribuir una amplitud concreta al sonido en cada instante. Por ejemplo, 8 bits ofrecen 256 niveles de cuantización y 16, 65536. El margen dinámico de la audición humana es de unos 100 dB. La división del eje se puede realizar a intervalos iguales o según una determinada función de densidad, buscando más resolución en ciertos tramos si la señal que se trata tiene más componentes en cierta zona de intensidad, como veremos en las técnicas de codificación.

El proceso completo se denomina habitualmente PCM (Pulse Code Modulation) y así nos referiremos a el en lo sucesivo.

2. Codificación y compresión.

Una significativa cantidad de la información que proporciona el PCM puede desecharse, en base al fenómeno conocido como enmascaramiento, que tiene relación con la percepción auditiva del oído humano. Así pues, lo importante no es gravar o reproducir fielmente el sonido original sin pérdida de todos sus matices, sino conseguir gravar y o reproducir los sonidos que realmente puede percibir el oído humano, ya que este no aprecia todas las frecuencias que pueden emitirse.

El proceso se realiza del siguiente modo: Al comprimir desechamos información que no va a ser percibida por el oído (algoritmo con pérdida) y con ello ahorramos espacio en el posterior almacenamiento. Cuando dos señales de frecuencia similar se superponen, sólo podemos escuchar la que posee el mayor volumen, efecto que se conoce como enmascaramiento o masking.

El oído humano percibe un rango de frecuencias entre 20 Hz. y 20 Khz. En primer lugar, la sensibilidad es mayor en la zona alrededor de los 2-4 Khz., de

forma que el sonido resulta más difícilmente audible cuanto más cercano a los extremos de la escala. En segundo lugar está el enmascaramiento, cuyas propiedades utilizan exhaustivamente los algoritmos más interesantes: cuando la componente a cierta frecuencia de una señal tiene una energía elevada, el oído no puede percibir componentes de menor energía en frecuencias cercanas, tanto inferiores como superiores. A una cierta distancia de la frecuencia enmascaradora, el efecto se reduce tanto que resulta despreciable; el rango de frecuencias en las que se produce el fenómeno se denomina banda crítica (critical band). Las componentes que pertenecen a la misma banda crítica se influyen mutuamente y no afectan ni se ven afectadas por las que aparecen fuera de ella. La amplitud de la banda crítica es diferente según la frecuencia en la que nos situemos y viene dada por unos determinados datos que demuestran que es mayor con la frecuencia. Hay que señalar que estos datos se obtienen por experimentos psicoacústicos, que se realizan con expertos entrenados en percepción sonora, dando origen con sus impresiones a los modelos psicoacústicos.

Este que hemos descrito es el llamado enmascaramiento simultáneo o en frecuencia. Existe, asimismo, el denominado enmascaramiento asimultáneo o en el tiempo, así como otros fenómenos de la audición que no resultan relevantes en este punto. En resumen, ciertas componentes en frecuencia de la señal admiten un mayor ruido del que generalmente consideraríamos tolerable y, por tanto, requieren menos bits para ser codificadas si se dota al codificador de los algoritmos adecuados para resolver máscaras.

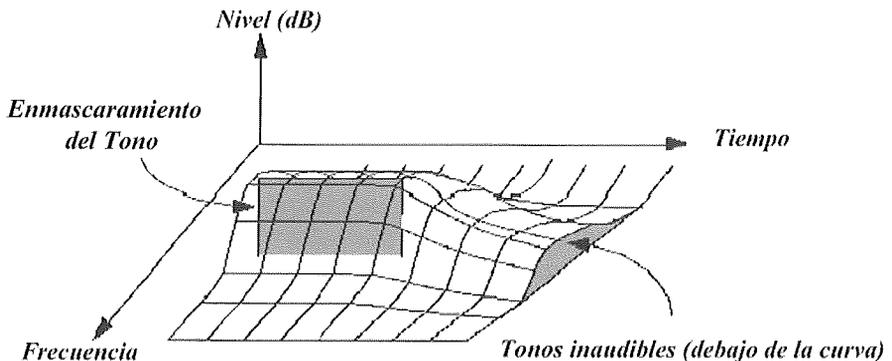


Figura 1: representación tridimensional del enmascaramiento.

La digitalización de la señal mediante PCM es la forma más simple de codificación de la señal, y es la que utilizan tanto los CD como los sistemas DAT. Toda digitalización, añade ruido a la señal, y cuantos menos bits se utilicen en el muestreo y la cuantización, mayor será el error al aceptar valores discretos para la señal continua, esto es, mayor será el ruido. Para evitar que el ruido alcance un nivel excesivo hay que emplear un gran número de bits, de forma que a 44'1 KHz. y utilizando 16 bits para cuantizar la señal, uno de los dos canales de un CD produce más de 700 kilobits por segundo (kbps). Como veremos, gran parte de esta información es innecesaria y ocupa un ancho de banda que podría liberarse, a costa de aumentar la complejidad del sistema decodificador e incurrir en cierta pérdida de calidad. El compromiso entre ancho de banda, complejidad y calidad es el que produce los diferentes estándares del mercado.

Calidad	Muestreo	Bits/muestra	Modo	Tasa de bits	Frecuencia
Teléfono	8 KHz	8	Mono	64 kbps	200-3400 Hz
Radio AM	11'025 KHz	8	Mono	88 kbps	
Radio FM	22'050 KHz	16	Estéreo	705'6 kbps	
CD	44'1 KHz	16	Estéreo	1411'2 kbps	20-20000 Hz
DAT	48 KHz	16	Estéreo	1536 kbps	20-20000 Hz

Tabla 1: comparación de formatos de calidad de audio

Un modo mejor de codificar la señal es mediante PCM no-lineal o cuantización logarítmica, que consiste en dividir el eje de la amplitud de tal forma que los escalones sean mayores cuanto más energía tiene la señal, con lo que se consigue una relación señal/ruido igual o mejor con menos bits. Con este método se puede reducir el canal de CD audio a 350 kbps, lo cual evidentemente es una mejora sustancial, aunque puede reducirse mucho más. Otros sistemas similares nos llevan a la cuantización adaptativa (APCM), diferencial (DPCM) y la mezcla de ambas, ADPCM. Así prosigue la reducción del ancho de banda, pero sin llegar a los niveles que proporciona el tener en cuenta los efectos del enmascaramiento.

El estándar MPEG Audio es el ejemplo más popular de SBC o codificación sub-banda (sbc: sub-band coding) que a diferencia de los métodos específicos para ciertas fuentes puede codificar cualquier señal de audio sin importar su origen, ya sea voz, música o sonido de tipos variados.

El sistema de compresión usa un proceso en el cual la señal se descompone en subbandas que crean a su vez otras subbandas, las cuales son comparadas con el modelo original para determinar cuales bandas son las que deben de utilizarse y cuales deben desecharse.

El principio básico del SBC es la limitación del ancho de banda, descartando información en frecuencias enmascaradas. El resultado simplemente no es el mismo que el original, pero si el proceso se realiza correctamente, el oído humano no percibe la diferencia.

La mayoría de los codificadores SBC utilizan el mismo esquema: un filtro o un banco de ellos, o algún otro mecanismo descompone la señal de entrada en varias subbandas. A continuación se aplica un modelo psicoacústico que analiza tanto las bandas como la señal y determina los niveles de enmascaramiento utilizando los datos psicoacústicos de que dispone. Considerando estos niveles de enmascaramiento se cuantizan y codifican las muestras de cada banda: si en una frecuencia dentro de la banda hay una componente por debajo de dicho nivel, se desecha. Si lo supera, se calculan los bits necesarios para cuantizarla y se codifica. Por último se agrupan los datos según el estándar correspondiente que estén utilizando codificador y decodificador, de manera que éste pueda descifrar los bits que le llegan de aquél y recomponer la señal.

La decodificación es mucho más sencilla, ya que no hay que aplicar ningún modelo psicoacústico. Simplemente se analizan los datos y se recomponen las bandas y sus muestras correspondientes.

En los últimos diez años la mayoría de las principales compañías de la industria del audio han desarrollado sistemas SBC. A finales de los años ochenta (1987), un grupo de estandarización del ISO llamado Motion Picture Experts Group (MPEG) junto con el Instituto FRAUNHOFER comenzó a desarrollar los estándares para la codificación tanto de audio como de vídeo en la norma ISO-MPEG Audio Layer (capa) 3 agrupados en ISO / IEC IS11172 - 3 y en ES 13818 - 3. Con ello se creó un sistema de compresión de audio que permitía obtener una calidad comparable a la de un CD pero necesitaba de ordenadores potentes para realizar la compresión. Solo en los últimos años han aparecido ordenadores domésticos con la suficiente potencia y memoria para realizar todo el proceso.

3. El estándar MPEG Audio

El estándar MPEG Audio contempla tres niveles diferentes de codificación-decodificación de la señal de audio, de los cuales sólo el primero está totalmente terminado. Los otros dos son aplicables, y de hecho se utilizan habitualmente, pero siguen abiertos a ampliaciones. Estos tres niveles son:

MPEG-1: «Codificación de imágenes en movimiento y audio asociado para medios de almacenamiento digital hasta 1'5 Mbit/s»

MPEG-2: «Codificación genérica de imágenes en movimiento e información de audio asociada»

MPEG-3: la planificación original contemplaba su aplicación a sistemas HDTV; finalmente fue incluido dentro de MPEG-2.

MPEG-4: «Codificación de objetos audiovisuales»

En el Apéndice 1 se detallan todos los niveles del estándar MPEG y las referencias ISO correspondientes.

A su vez, MPEG describe tres esquemas de codificación de audio denominados esquema-1, esquema-2 y esquema-3. Del primero al tercero aumentan tanto la complejidad del codificador como la calidad del sonido. Los tres son compatibles jerárquicamente, esto es, el decodificador esquema-1 es capaz de interpretar información producida por un codificador esquema-1 y todos los niveles por debajo del 1. Así, un decodificador esquema-3 acepta los tres niveles de codificación, mientras el esquema-2 sólo acepta el 1 y el 2.

MPEG define, para cada esquema, el formato del bitstream y el decodificador (que puede ser implementado de diferentes maneras). Con vistas a admitir futuras mejoras no se define el codificador, pero en un apartado informativo se da un ejemplo de codificador para cada uno de los esquemas. Hay que decir que tanto MPEG-1 como MPEG-2 emplean estos tres esquemas, pero este último añade nuevas características

En lo referente al MPEG-1 el sistema que describe la norma ISO es:

- Codificación: el codificador procesa la señal de audio digital y produce el bitstream empaquetado para su almacenamiento y/o transmisión. El algoritmo de codificación no está determinado, y puede utilizar enmascaramiento, cuantización variable y escalado. Sin embargo, debe ajustarse a las especificaciones del decodificador.

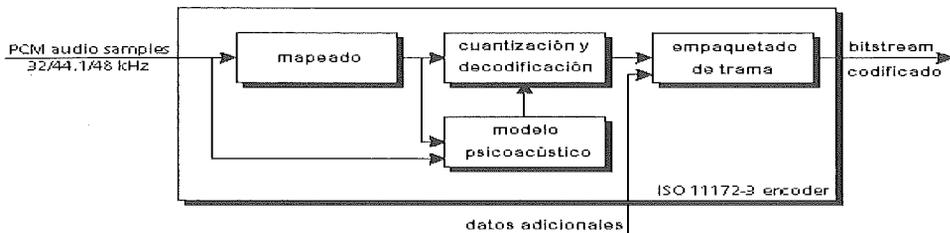
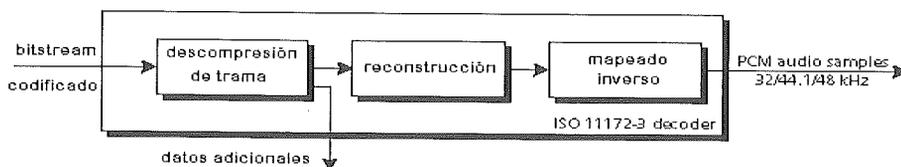


Figura 2: codificador según la norma ISO 11172-3

Las muestras se introducen en el codificador y a continuación el mapeador crea una representación filtrada y submuestreada de la señal de entrada. Las muestras mapeadas se denominan tanto muestras de subbanda (esquemas 1 y 2) como muestras de subbanda transformadas (esquema 3). El modelo psicoacústico crea una serie de datos (dependiendo de la implementación del codificador) que sirven para controlar la cuantización y codificación. Este último bloque crea a su vez su propia serie de datos, de nuevo dependiendo de la implementación. Por último, el bloque de empaquetamiento de trama se encarga de agrupar como corresponde todos los datos, pudiendo añadir algunos más, llamados datos adicionales, como por ejemplo CRC o información del usuario.



- Decodificación: el decodificador debe procesar el bitstream para reconstruir la señal de audio digital. La especificación de este elemento sí está totalmente definida y debe seguirse en todos sus puntos. La figura ilustra el esquema del decodificador.

Los datos del bitstream son desempaquetados para recuperar las diversas partes de la información. El bloque de reconstrucción recompone la versión cuantizada de la serie de muestras mapeadas. El mapeador inverso transforma estas muestras de nuevo a PCM.

- Esquemas:

1. Incluye la división del mapeado básico de la señal de audio digital en 32 subbandas, segmentación para el formateo de los datos, modelo psicoacústico y cuantización fija. El retraso mínimo teórico es de 19 ms.
2. Incluye codificación adicional, factores de escala y diferente composición de trama. El retraso mínimo teórico es de 35 ms.
3. Incluye incremento de la resolución en frecuencia, basado en el uso de un banco de filtros híbrido. Cuantización no uniforme, segmentación adaptativa y codificación entrópica de los valores cuantizados. El retraso mínimo teórico es de 59 ms.

Esquema	Objetivo	Compresión	Calidad 64 kbps	Calidad 128 kbps	Retardo
Esquema-1	192 kbps	4 a 1			19 ms
Esquema-2	128 kbps	6 a 1	2'1 a 2'6	Más de 4	35 ms
Esquema-3	64 kbps	12 a 1	3'6 a 3'8	Más de 4	59 ms

Tabla 2: resumen de datos de los tres esquemas

La calidad viene dada del 1 al 5, siendo el 5 la superior (ver apartado 6). Hay que señalar que pese a los números de la norma ISO, el retraso típico acostumbra a ser tres veces mayor en la práctica.

- Modos: hay cuatro modos de funcionamiento para cualquiera de estos tres esquemas.
 - a. single channel o canal único: una señal en un bitstream.
 - b. dual channel o canal doble: dos señales independientes en un mismo bitstream.
 - c. estéreo: como el anterior, perteneciendo las señales al canal izquierdo y derecho de una señal estéreo original.
 - d. joint stereo: como el anterior, aprovechando ciertas características del estéreo como irrelevancia y redundancia de datos para reducir la tasa de bits.

3.1. Funcionamiento del sistema.

Las características y diferencias entre los tres esquemas de MPEG-1 son: .

a. La codificación:

a.1) El banco de filtros: realiza el mapeado del dominio del tiempo al de la frecuencia. Existen dos tipos: el polifase y el híbrido polifase/MDCT. Estos bancos proporcionan tanto la separación en frecuencia para el codificador como los filtros de reconstrucción del decodificador. Las muestras de salida del banco están cuantizadas.

a.2) El modelo psicoacústico: calcula el nivel a partir del cual el ruido comienza a ser perceptible, para cada banda. Este nivel se utiliza en el bloque de asignación de bit/ruido para determinar la cuantización y sus niveles. De nuevo, se utilizan dos diferentes. En ambos, los datos de salida forman el SMR (signal-to-mask ratio) para cada banda o grupo de bandas.

a.3) Asignación de bit/ruido: examina tanto las muestras de salida del banco de filtros como el SMR proporcionado por el modelo psicoacústico, y ajusta la asignación de bit o ruido, según el esquema utilizado, para satisfacer simultáneamente los requisitos de tasa de bits y de enmascaramiento.

a.4) El formateador de bitstream: toma las muestras cuantizadas del banco de filtros, junto a los datos de asignación de bit/ruido y otra información lateral para formar la trama.

Los tres esquemas utilizan diferentes algoritmos para cumplir estas especificaciones:

Esquema I:

- El mapeado tiempo-frecuencia se realiza con un banco de filtros polifase con 32 subbandas. Los filtros polifase consisten en un conjunto de filtros con el mismo ancho de banda con interrelaciones de fase especiales que ofrecen una implementación eficiente del filtro subbanda. Se denomina filtro subbanda al que cubre todo el rango de frecuencias deseado. En general, los filtros polifase combinan una baja complejidad de computación con un diseño flexible y múltiples opciones de implementación.
- El modelo psicoacústico utiliza una FFT (Fast Fourier Transform) de 512 puntos para obtener información espectral detallada de la señal. El resultado de la aplicación de la FFT se utiliza para determinar los enmascaramientos en la señal, cada uno de los cuales produce un nivel de enmascaramiento, según la frecuencia, intensidad y tono. Para cada subbanda, los niveles individuales se combinan y forman uno global, que se compara con el máximo nivel de señal en la banda, produciendo el SMR que se introduce en el cuantizador.
- El bloque de cuantización y codificación examina las muestras de cada subbanda, encuentra el máximo valor absoluto y lo cuantiza con 6 bits. Este valor es el factor de escala de la subbanda. A continuación se determina la asignación de bits para cada subbanda minimizando el NMR (noise-to-mask ratio) total. Es posible que algunas subbandas con un gran enmascaramiento terminen con cero bits, es decir, no se codificará ninguna muestra. Por último las muestras de subbanda se cuantizan linealmente según el número de bits asignados a dicha subbanda concreta.
- El trabajo del empaquetador de trama es sencillo. La trama, según la definición ISO, es la menor parte del bitstream decodificable por sí misma. Cada trama empieza con una cabecera para sincronización y diferenciación, así como 16 bits opcionales de CRC para detección y corrección de errores. Se emplean, para cada subbanda, 4 bits para describir la asignación de bits y otros 6 para el factor de escala. El resto de bits en la trama se utilizan para la información de samples, 384 en total, y con la opción de añadir cierta información adicional. A 48 KHz, cada trama lleva 8 ms de sonido.

Esquema II:

- El mapeado de tiempo-frecuencia es idéntico al del esquema I.
- El modelo psicoacústico es similar, salvo que utiliza una FFT de 1024 puntos para obtener mayor resolución espectral. En los demás aspectos, es idéntico.
- El bloque de cuantización y codificación también es similar, generando factores de escala de 6 bits para cada subbanda. Sin embargo, las tramas del esquema II son tres veces más largas que las del esquema I, de forma que se concede a cada subbanda tres factores de escala, y el codificador utiliza uno, dos o los tres, según la diferencia que haya entre ellos. La asignación de bits es similar a la del esquema I.- El formateador de trama: la definición ISO de trama es la misma que en el punto anterior. Utiliza la misma cabecera y estructura de CRC que el esquema I. El número de bits que utilizan para describir la asignación de bits varía con las subbandas: 4 bits para las inferiores, 3 para las medias y dos para las superiores, adecuándose a las bandas críticas.

Los factores de escala se codifican junto a un número de dos bits que indica si se utilizan uno, dos o los tres. Las muestras de subbanda se cuantizan y a continuación se asocian en grupos de tres, llamados gránulos. Cada uno se codifica con una palabra clave, lo que permite interceptar mucha más información redundante que en el esquema I. Cada trama contiene, pues, 1152 muestras PCM. A 48 KHz. cada trama lleva 24 ms de sonido.

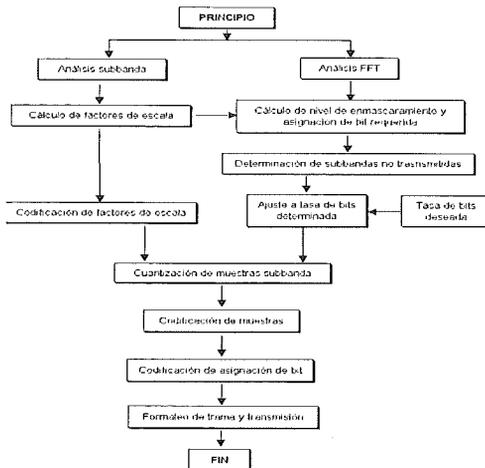


Diagrama 1: diagrama de flujos del codificador para esquema-1 y esquema-2 según ISO 11172-3

Esquema III:

El esquema III es sustancialmente más complicado que los dos anteriores e incluye una serie de mejoras cuyo análisis resultaría desbordante, de manera que no entraremos en tantos detalles. Su diagrama de flujos es conceptualmente semejante al visto para los otros dos esquemas, salvo que se realizan múltiples iteraciones para procesar los datos con el mayor nivel de calidad en un cierto tiempo, lo cual complica su diseño hasta el punto de que los diagramas ISO ocupan decenas de páginas.

- El mapeado de tiempo-frecuencia añade un nuevo banco de filtros, el DCT (Discrete Cosine Transform), que con el polifase forman el denominado filtro híbrido. Proporciona una resolución en frecuencia variable, 6x32 o 18x32 subbandas, ajustándose mucho mejor a las bandas críticas de las diferentes frecuencias.
- El modelo psicoacústico es una modificación del empleado en el esquema II, y utiliza un método denominado predicción polinómica. Incluye los efectos del enmascaramiento temporal.
- El bloque de cuantización y codificación también emplea algoritmos muy sofisticados que permiten tramas de longitud variable. La gran diferencia con los otros dos esquemas es que la variable controlada es el ruido, a través de bucles iterativos que lo reducen al mínimo posible en cada paso.
- El formateador de trama: la definición de trama para este esquema según ISO varía respecto de la de los niveles anteriores: «mínima parte del bitstream decodificable mediante el uso de información principal adquirida previamente». Las tramas contienen información de 1152 muestras y empiezan con la misma cabecera de sincronización y diferenciación, pero la información perteneciente a una misma trama no se encuentra generalmente entre dos cabeceras.

La longitud de la trama puede variarse en caso de necesidad. Además de tratar con esta información, el esquema III incluye codificación Huffman de longitud variable, un método de codificación entrópica que sin pérdida de información elimina redundancia. Los métodos de longitud variable se caracterizan, en general, por asignar palabras cortas a los eventos más frecuentes, dejando las largas para los más infrecuentes.

b) La decodificación:

Es mucho más sencilla que la codificación, de manera que con lo ya comentado en partes anteriores basta para seguir los siguientes diagramas ISO que incluyen algunas notas aclaratorias al margen que no forman parte de las figuras originales de la norma.

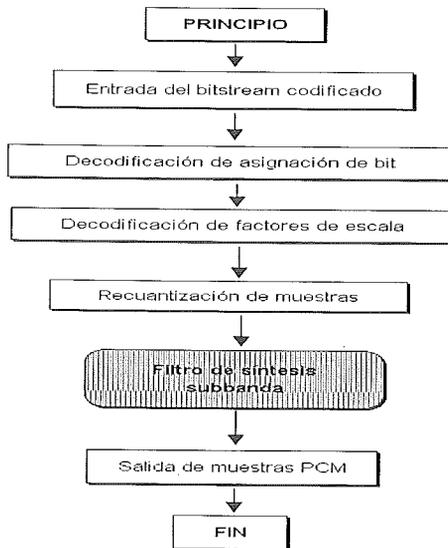


Diagrama 2: diagrama de flujos del decodificador para esquema-1 y esquema-2 según ISO 11172-3

El esquema-1 obtiene la mayor calidad de sonido a 384 kbps. Las aplicaciones para las que resulta más útil son las relacionadas con la grabación, tanto en cinta como disco duro o discos magneto-ópticos, que aceptan esta tasa de bits sin problemas.

El esquema-2 produce sus mejores resultados de calidad a 256 kbps, pero se mantiene en un nivel aceptable hasta los 64 kbps. Esto hace que se utilice en transmisión de audio, televisión, grabación profesional o doméstica y productos multimedia.

El esquema-3, para una determinada calidad de sonido ofrece la menor tasa de bits y viceversa, fijando la tasa de bits ofrece la mejor calidad posible.

El esquema-3 está orientado a aplicaciones donde la necesidad de un ancho de banda reducido justifique el costoso y sofisticado sistema de codificación. La calidad es excelente hasta 64 kbps, de forma que se utiliza en telecomunicaciones y sistemas de sonido profesional, así como al nivel de usuario por parte de aficionados con formación informática.

4. El uso del MPEG en la industria audiovisual.

En la industria y el mundo del audio se viene ampliando su uso para Conexiones musicales via RDSI.

Las redes telefónicas digitales (RDSI=Red Digital de Servicios Integrados) ofrecen servicios seguros de conexión con dos canales de datos de 64 kbps por adaptador; en otras redes los canales son de 56 kbps, pero en ambas los costes de transmisión son similares a las líneas telefónicas tradicionales, analógicas, que permiten un máximo de 33'6 ó 56 kbps (vía módem). Con el esquema-3 una conexión de banda estrecha RDSI de bajo coste permite transmitir sonido con calidad CD. Los estudios de sonido y estaciones de transmisión se benefician de la posibilidad de la «música por teléfono» de varias maneras. Se ahorra dinero, pues sólo se paga el tiempo de transmisión, a diferencia de la línea telefónica y únicamente se emplea un pequeño conector RDSI para cada canal. Los programas pueden aumentar su atractivo, ofreciendo tomas de alta calidad y noticias en directo sin la pérdida de calidad del sonido telefónico. Aparecen nuevos campos, como el Estudio Virtual, donde artistas en distintas localidades pueden tocar y grabar juntos sin necesidad de viajar hasta el estudio en sí.

Ya en 1992, Radio FFN, una estación privada de Niedersachsen, Alemania, reemplazó sus líneas telefónicas tradicionales por conexiones ISDN y codecs esquema-3, para transmitir 8 programas locales diarios al estudio central. El ahorro declarado en cuotas de transmisión fue de 300.000 \$ anuales.

- En uno de los primeros ensayos reales de la potencia de este sistema, todas las estaciones de radio privadas en Alemania utilizaron codecs esquema-3 durante los Juegos Olímpicos de Invierno en Albertville para conectar las diferentes localizaciones del evento con la sede central en Meribel, con gran éxito.

5. - Transmisión digital por satélite.

Actualmente se encuentra en plena expansión y se está construyendo un sistema de transmisión de sonido digital a escala mundial por satélite. El nombre del proyecto es Worldstar y utiliza satélites en órbita geoestacionaria. Cada uno está equipado con tres canales de conexión que se pueden multiplexar en hasta 96 subcanales de 16 kbps. Estos son combinables dinámicamente, de manera que se pueden agrupar para formar canales de hasta 128 kbps de capacidad, codificados con el esquema-3. Así, se pueden utilizar cuatro subcanales para formar uno de 64 kbps para transmitir un concierto y al finalizar, utilizar cada uno de ellos para enviar las noticias en cuatro idiomas diferentes.

6. - Audio en Internet.

Internet es una red mundial de conmutación de paquetes con cientos de miles de máquinas unidas entre sí por medio de varios sistemas de comunicaciones. Los proveedores profesionales normalmente acceden a la red a través de enlaces con un ancho de banda muy elevado (hasta 2 Gbps). El consumidor doméstico, sin embargo, utiliza canales de bajo coste y ancho de banda limitado (RDSI de 64 kbps o conexión telefónica de 28'8 - 56 kbps). La tasa de transmisión efectiva varía en función del uso de la parte de la red accedida, situándose en algún punto entre cero y la máxima capacidad del módem.

Sin la codificación de audio, obtener ficheros de audio sin comprimir de un servidor remoto llevaría a unos tiempos de transmisión excesivamente largo. Por ejemplo, suponiendo que se alcanza la tasa de 28'8 kbps (lo cual es bastante optimista) una pista de CD de sólo tres minutos tardaría más de dos horas en recibirse. Por tanto, el audio en Internet exige un método de codificación que ofrezca la mayor calidad posible a la vez que permita la decodificación en tiempo real para un amplio número de plataformas sin necesidad de ampliar el hardware, aunque incluya esta posibilidad como elemento de hipotéticas tarjetas de ampliación. Por supuesto, la elección es el esquema-3. Hay varios reproductores en tiempo real, como el Winplay3, AudioCatalyst, Fraunhofer MP3 Encoder, MusicMatch, Jukebox, XingMP3 Encoder o BladeEncoder que ofrecen el servicio requerido.

En la práctica, las expectativas se han cumplido con creces, de tal manera que el fenómeno MP3 está en plena expansión en la telaraña mundial. Ya hay innumerables servidores que ofrecen música en el formato esquema-3 (ficheros de extensión. MP3), de los cuales forman parte tanto aficionados como casas de grabación y grupos independientes. Además, se incorporan temas en este formato a las páginas WEB como elemento para incrementar su atractivo, de forma similar a como se venía haciendo con el MIDI, salvando la barrera de las muy inferiores posibilidades de este.

Llegados a este punto, hay que señalar la importancia de respetar los copyrights a la hora de incluir temas de música en un servidor, así como al almacenarlos en disco duro o CD-ROM. La perspectiva de duplicar la capacidad de los CDs tradicionales resulta sumamente interesante a la comunidad informática, y dado el auge de las grabadoras domésticas puede decirse que el mercado pirata de CDs conteniendo las discografías al completo de diversos grupos o compositores es una realidad, sea con ánimo de lucro o no. El más que previsible auge del DVD-ROM como estándar en el futuro cercano no supone sino un agravamiento del problema.

Las aplicaciones legales que se conocen hasta ahora son, por ejemplo, las de Opticom, Cerberus Sound y Madison. La primera ofrece soluciones para que las

casas ofrezcan a los clientes audio por demanda, enviando los temas seleccionados al ordenador remoto del usuario. Cerberus se dedica a la comercialización directa de estos temas como un sistema más de venta electrónica. Madison se asegura el pago de los derechos de reproducción y cuya distribución se realiza en línea a través del denominado SDMI (Segure Digital Music Initiative). Asimismo se avanza en el concepto de Internet Radio, dado que se obtiene calidad superior a la de la onda corta con un ancho de banda tan escaso como 16 kbps. Opticom de nuevo se puso a la cabeza en este campo, junto a Telos, compañía que asociada con Apple presentó en Septiembre del 96 la tecnología Audioactive. Otras compañías como Microsoft, a través de su tecnología multimedia Netserver, e IBM entre otras han presentado sus alternativas como el Sistema de encriptación Electronic Music Management System (EMMS), el Software Liquid Audio o el MS 4.0 Audio que tratan de ofrecen sistemas para salvaguardar los derechos de autor y los de las compañías discográficas.

Apéndice 1: especificaciones ISO MPEG.

A continuación se refieren las diferentes partes que componen el estándar MPEG-1 y MPEG-2.

MPEG-1: Cinco partes, las tres primeras estandarizadas desde 1992.

IS-11172-1 (Sistema) - Describe la sincronización y multiplexación de señales de audio y vídeo.

IS-11172-2 (Vídeo) - Describe la compresión de señales de vídeo, centrándose en el escaneo progresivo y considerando especialmente las aplicaciones de vídeo en CD.

IS-11172-3 (Audio) - Describe una familia genérica de codificación de audio, con tres miembros jerárquicamente compatibles, denominados esquema-1, esquema-2 y esquema-3.

IS-11172-4 (Tests de conformidad) - Describe los procedimientos para determinar las características de los bitstreams codificados y el proceso de decodificación, así como los tests de conformidad con los requerimientos establecidos en las otras partes.

DTR-11172-5 (Simulación por software) - Es un informe técnico sobre la implementación por software de las tres primeras partes de MPEG-1.

MPEG-2: Nueve partes, las tres primeras estandarizadas desde 1994, con algunos añadidos posteriores. *En diferentes estados de acabado.*

IS-13818-1 (Sistema) - Describe la sincronización y multiplexación de señales de audio y vídeo; estandarizado por ITU-T como H.222.

IS-13818-2 (Vídeo) - Describe un conjunto genérico de herramientas para codificación de vídeo; estandarizado por ITU-T como H.262.

IS-13818-3 (Audio) - Describe una extensión de MPEG-1 compatible hacia atrás, para codificación de audio multicanal (sonido envolvente, sonido multilingüe) y una extensión no compatible hacia atrás para frecuencias de muestreo inferiores, para soportar aplicaciones de sonido con requerimientos de ancho de banda limitado.

IS-13818-4 (Tests de concordancia) - Describe los procedimientos para determinar las características de los bitstreams codificados y el proceso de decodificación, así como los tests de conformidad con los requerimientos establecidos en las otras partes.

DTR-13818-5 (Simulación por software) - Es un informe técnico sobre la implementación por software de las tres primeras partes de MPEG-2.

IS-13818-6 (Extensiones de sistema - Control y comandos para medios de almacenamiento digital) - describe un conjunto de protocolos para aplicaciones cliente-servidor.

CD-13818-7 (Audio, codificación no compatible hacia atrás (NBC)) - Describe un esquema de codificación de audio mejorado para señales mono y estéreo, así como para sonido multicanal.

13818-8 (Vídeo, extensión para muestras de entrada de 10 bits) - Se ha retirado, debido al escaso interés.

IS-13818-9 (Especificación del interface en tiempo real para aplicaciones low-jitter) - Define las restricciones temporales para el envío en tiempo real de bitstreams MPEG-2.

WD-13818-10 (Extensiones de concordancia - DSM-CC) - Describe los añadidos a IS-13818-4 para DSM-CC.

Apéndice 2: *el modelo psicoacústico.*

Los modelos psicoacústicos se componen a partir de las percepciones de un grupo de personas entrenadas para rendir al máximo en este campo. Por medio de una serie de experimentos se puede determinar la sensibilidad del oído humano a una serie de fenómenos, de forma que aparezcan resultados útiles para el tratamiento del sonido, como ya hemos visto.

Para cuantificar la Sensibilidad del oído humano se puede realizar el experimento de situar a una persona aislada de otros sonidos y ofrecer un tono de 1 Khz al nivel mínimo de sonido posible e ir elevando el volumen hasta que sea justo perceptible. (Figura 4)

UMBRAL DE SILENCIO

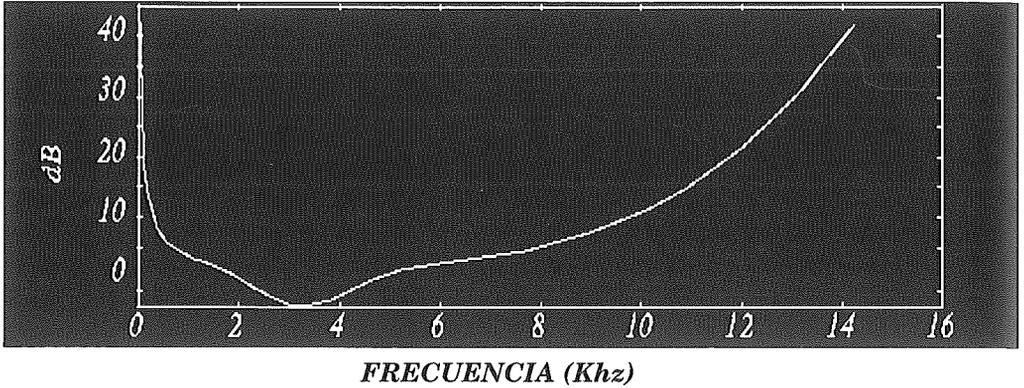


Figura 4: sensibilidad del oído humano en función de la frecuencia

Para cuantificar el Enmascaramiento en frecuencia se experimenta en las mismas condiciones: ofrecer un tono de 1 KHz (enmascarador) a un volumen determinado (60 dB, por ejemplo). Ofrecer un sonido de prueba a 1'1 KHz y elevar su volumen hasta que sea justo perceptible. (Figura 5)

ENMASCARAMIENTO DE TONO EN 1 kHz

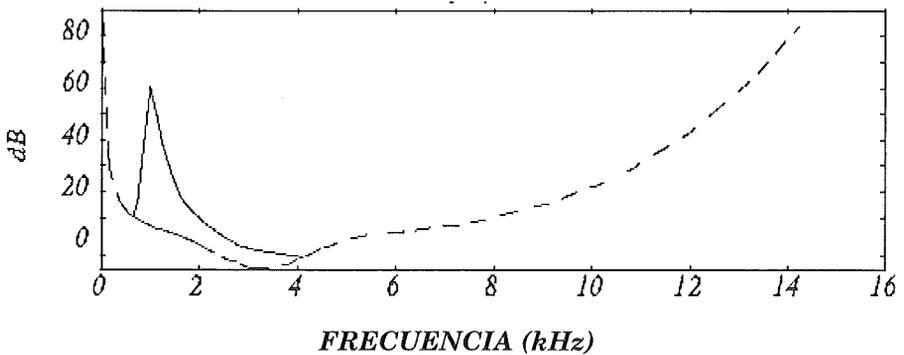


Figura 5: enmascaramiento en frecuencia del tono de 1 KHz

Otro tipo de experimentación es comprobar el enmascaramiento de diversos tonos de frecuencia que se realiza mediante la repetición del experimento anterior (figura 5) pero para todas las frecuencias de tonos de prueba necesarias. (Figura 6).

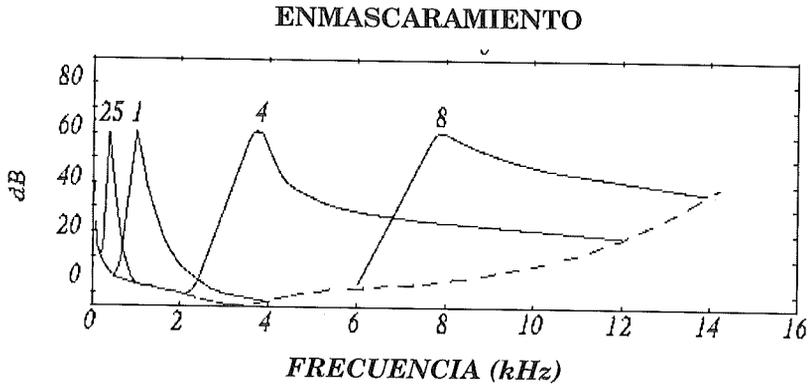


Figura 6: enmascaramiento de diversos tonos de prueba

Para poder comprobar el enmascaramiento temporal se realizará el experimento en las mismas condiciones que los anteriores, ofrecer un tono enmascarador de 1 KHz a 60 dB y un tono de prueba de 1'1 KHz a 40 dB. El tono de prueba no es audible. Parar el tono enmascarador y, con un pequeño retraso, el de prueba. Aumentar el retraso hasta que se distinga el tono de prueba. (Figura 7)

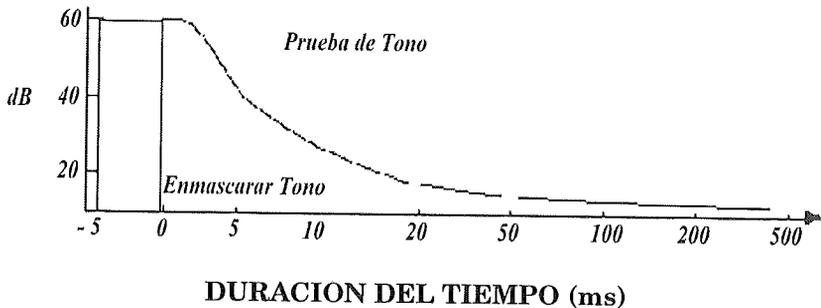


Figura 7: variación del volumen audible con el retardo

CONCLUSIÓN.

La principal característica de los formatos basados en los estándares MPEG es que son de dominio público. La Organización Internacional de Estándares (ISO) ha adoptado estos formatos como referencia, de forma que todo fabricante de software o hardware pueda adaptarse a ellos sin necesidad de pagar derechos a ninguna compañía privada. Las imágenes JPG, los archivos MP3, AVI, MPG, etc. que descargamos de Internet llevan en su seno la impronta del grupo de investigadores que formaron el MPEG: 'Motion Picture Expert Group' (Grupo de Expertos de Imágenes en Movimiento). Sus desarrollos están presentes hoy en día en cada imagen, vídeo o audio que exista en Internet, TV, radio y cine.

El mundo multimedia les debe tanto que casi se puede afirmar que sin ellos no existiría (o sería un caos).

Sin embargo, estos formatos no están quietos. Sólo las partes más elementales de ellos están fijas, de forma que son estándares capaces de evolucionar en el tiempo sin perder el hilo conductor que los define.

De esta forma, la familia de estándares MPEG sirve tanto para transmitir voz digital por Internet como para emitir vídeo, imagen fija o música de alta calidad. Todo es lo mismo. Un sistema de capas (o envoltorios) de información hace que los contenidos puedan ser diferentes sin perder el nexo común básico. Las capas más internas son absolutamente iguales para todos los formatos, mientras que las externas permiten flexibilidad propia para cada tipo de información multimedia.

El MP3 es un formato de compresión de audio que consigue una ratio de compresión de entre 1:11 y 12 sin pérdida de calidad apreciable. Esto significa que en un disco compacto se pueden grabar unos 11 CD's-Audio (aprox. unas 150 - 160 canciones).

El único problema es que para poder escuchar las canciones tendrá que hacerlo a través del ordenador (en la actualidad o mediante equipos que se acoplan a la cadena de música todavía costosos), con un programa reproductor de MP3. Además está el problema de la legalidad, es legal que se hagan copias como sistema de seguridad, si tiene los originales, pero es ilegal bajar o intercambiar MP3 desde Internet, a no ser que pague los derechos de autor.

Para obtener un archivo MP3 con calidad CD se deben hacer dos cosas. Primero debe extraerse la información del CD-Audio y pasarla al disco duro. Esto lo hace un programa denominado extractor (ripper o ripeador). De esta manera ya tenemos la información del CD-Audio en el disco duro, generalmente en un voluminoso fichero WAV.

Segundo, debido al gran tamaño de estos ficheros, necesitamos comprimirlos, y aquí es donde surge el compresor, que nos pasará este fichero WAV a un

fichero MP3, el cual contiene prácticamente la misma calidad de sonido pero ocupa 11 veces menos espacio.

Pie para Codificación El codificador procesa la señal de audio digital y produce el bitstream empaquetado para su almacenamiento y/o transmisión. El algoritmo de codificación puede utilizar enmascaramiento, cuantización variable y escalado. Las muestras se introducen en el codificador y a continuación el mapeador crea una representación filtrada y submuestreada de la señal de entrada. El modelo psicoacústico crea una serie de datos que sirven para controlar la cuantización y codificación. El bloque de empaquetamiento de trama se encarga de agrupar como corresponde todos los datos, pudiendo añadir algunos datos adicionales.

Finalmente, para poder escuchar la canción necesitaremos un reproductor de MP3, como el WinAmp. El winAMP es el reproductor de archivos de sonido del tipo MP3 más extendido. Para ser capaz de reproducir este tipo de ficheros se deberá tener como mínimo un ordenador con un procesador 486 a 100 MHz con una tarjeta de sonido de 16 bits. Para instalarlo simplemente se tiene que ejecutarlo y éste se guardará en el directorio asignado, asociando la extensión MP3 con él mismo, para que al hacer doble click sobre un archivo MP3 éste se ejecute automáticamente.

Pie para Decodificación El decodificador debe procesar el bitstream para reconstruir la señal de audio digital. Los datos del bitstream son desempaquetados para recuperar las diversas partes de la información. El bloque de reconstrucción recompone la versión cuantizada de la serie de muestras mapeadas. El mapeador inverso transforma estas muestras de nuevo a PCM (Pulse Code Modulation).

Cuadro: Las capas de MPEG1

Capa	Compresión	Transferencia
1	4:1	384 kbits/s
2	6:1 a 8:1	256 a 192 kbits/s
3	10:1 a 12:1	128 a 112 kbits/s

Respecto a la calidad del MP3, puede ser como uno quiera, menor calidad menor tamaño y menos tiempo empleado, pero para obtener un sonido con calidad digital la tasa de transferencia de bits o bitRate deberá ser de 128 Kbits/seg. con una frecuencia de muestreo de 44.100 KHz. Esto hay que tenerlo en cuenta a la hora de crear el MP3, pero por lo general es la opción que viene como estándar.

Resumiendo, hemos conseguido tener la canción que nos gusta en un fichero MP3 que nos ocupa unos 4 Megas de espacio. Y aquí es donde vienen los problemas legales, los derechos de autor.

Tanto intercambiar como bajar MP3s desde Internet es ilegal, a no ser que se paguen derechos de autor. El problema para las discográficas (no tanto para los autores debido al estrecho margen de beneficios que se llevan) es que existen multitud de sitios en los que pueden descargar MP3 y además puede decirse que existe impunidad en la red, lo único que se hace es hablar con el webmaster o administrador de la página, y normalmente o elimina los MP3 o los cambia a servidores extranjeros. De esta manera el único problema que se puede encontrar es que el enlace al MP3 no exista o que la velocidad de descarga sea desesperante.

Últimamente las casas discográficas están empezando a preocuparse por la futura industria legal de MP3, y prueba de ello son las llamadas arañas que se están desarrollando. Estas arañas son sistemas de búsquedas de canciones que localizan posibles canciones y comprueban si son legales. En muchos sitios encontrarás canciones MP3 pero con extensión cambiada, por ejemplo a zip o cualquier otra. Esto se hace para que existan menos posibilidades de localizar la canción, ya sea por las discográficas o por los administradores del sitio cuando este es público y gratuito, como Geocities, Xoom, Angelfire, y otros

El futuro del MP3 es muy alentador, ya que cada día más empresas lanzan al mercado reproductores para poder oírlos sin necesidad de computadora, ya sea como complemento a una cadena de música, como una especie de walkman o como reproductor para portátil. Además, cada vez más grupos lanzan a la red sus canciones de manera gratuita en sus propias páginas oficiales, como manera de darse a conocer. No solamente esto, sino que en la red existen empresas que comercializan legalmente con archivos MP3, pagando derechos de autor, por lo que son totalmente legales. Además los precios son muy baratos, por lo que las discográficas están empezando a tomarse el tema en serio, ya que la gente empieza a hacerse Cds a medida, sin contar para ello con las discográficas

MP3 en los videojuegos.

El formato de compresión MP3 es utilizado comercialmente por varias compañías de videojuegos en sus nuevos productos, gracias a las posibilidades de introducir mucha más música y de semejante calidad a las pistas de audio habituales (como 3DO en la banda sonora de *Might and Magic III*, o *Sierra en Middle Earth* ambientado en el universo de Tolkien).

En la actualidad se está empezando a publicar y mucho sobre el MP4, mayor calidad de sonido y menor tamaño que el MP3, pero el tema todavía está algo verde, tanto que incluso una compañía comercial ha patentado un formato de audio que ha llamado para mayor confusión MP4, y a diferencia del MP3, los usuarios particulares no podrán hacerse sus MP4, por lo que de esta manera se evitará la piratería. Eso sí, las empresas que quieran comercializar canciones con este MP4 deberán pagar por la patente. Por todo esto el tema está bastante liado, y no hay mucha información sobre el formato sucesor del MP3, unos dicen que se llamará MP4, otros dicen que será muy similar al formato VQF (que es mejor que el MP3, pero no está tan extendido).

MPEG Layer 4. Existen dos versiones (1 y 2) en función del tipo de contenidos. La primera ya ha sido adoptada por el ISO, mientras la segunda está todavía en fase de experimentación.

Esta cuarta capa, sin duda, va a revolucionar el concepto que tenemos de Internet. A partir de este desarrollo la información va a ser mucho más compacta en función de su propio contenido, lo cual permitirá introducir animación 3D, secuencias de voz en off sobre vídeo o hiperenlaces entre diferentes medios de una manera más lógica.

Cada tipo de información será codificada en su formato natural, de forma que una secuencia de lo que hoy llamamos vídeo, puede estar hecha con animación 3D y voz en off, y sólo ocupará el espacio imprescindible (mucho menos que si se almacena la señal de vídeo directamente). Si la siguiente escena está captada del mundo real con una cámara, se codificará como vídeo tradicional comprimido (mucho menos compacto).

Otro ejemplo: la voz humana sólo es capaz de emitir frecuencias de hasta 8 KHz. ¿Para qué usar audio de mayor ancho de banda? Si un vídeo tiene voz en off es mejor codificarla como voz, no como sonido de alta calidad.

Otro más: el audio musical estéreo dispone de dos canales que, aunque se parecen mucho, tienen ligeras diferencias. Esto nos hace pensar que, en realidad, no es el doble que el audio en mono, sino una vez y media aproximadamente.

Los archivos MPEG ya no sólo serán un montón de información visual o sonora embutida 'a capón', sino que tendrán cierta 'inteligencia' adicional, para distinguir el contenido y tratar cada cual según el criterio óptimo.

AAC. Mas allá del MP3

El AAC es una especificación más moderna que forma parte del estándar MPEG-2 y cuyas siglas significan Codificación Avanzada de Audio. AAC es un algoritmo más eficaz que la capa 3 de MPEG-1, consiguiendo niveles de compresión mayores y además mejorando la calidad. Sin embargo no es compatible con los formatos anteriores, MPEG-1, y para evitar los problemas que esto puede suponer, también se ha definido MPEG-2 Multicanal o MPEG-2 BC (Backward Compatible). Mientras los MP3 requieren un ancho de banda de 128 Kb/s para conseguir calidad próxima a la del CD, con AAC tan sólo son necesario 96 para conseguir mejor calidad. Ya se ha aplicado en un servicio de descarga de música de A.T.T. (utilizando ficheros encriptados para preservar los derechos del autor). Como AAC es el sucesor de MP3 se ha dado en llamar MP4, lo que da lugar a confusiones y a hablar erróneamente de MPEG-4.

Enlaces hacia sitios sobre MP3

A continuación los principales sitios relacionados con MP3:

- **MP3.com:** Aquí se encuentra cientos de canciones legales de MP3 y mucha información sobre este formato musical.
- **Todo sobre MP3:** Un sitio informativo y muy interesante sobre MP3 que vale la pena revisar por sus excelentes contenidos
- **Buscador de MP3 en Lycos :** Por medio de Lycos se puede encontrar una gran cantidad de MP3 en Internet
- **Goodnoise:** Dedicado a promover el MP3 y a ayudar al navegante en la recolección de éstos.
- **MusicMatch:** Un sitio donde se encuentra música en MP3 y software para reproducir y comprimir este tipo de música.
- **WinAmp:** El reproductor de MP3s más popular y de distribución gratuita
- **Sonique:** Un nuevo reproductor que ha despertado la atención por su facilidad de uso y de distribución gratuita

- **Override:** Una opción bastante buena y en español para oír música MP3
- **AudioCatalyst 1.5:** Un buen compresor de música para poder crear música en formato MP3
- **Electronic Cosmo's MPEG Suite 98:** Otro compresor muy popular para hacer nuestros propios MP3

Referencias.

- **Código C++ informativo:** <ftp://ftp.fhg.de/pub/layer3>
ftp://ftp.tnt.uni-hannover.de/pub/MPEG/audio/mpeg2/public_software
- **Cerberus Central Ltd.** <http://www.cdj.co.uk/>
- **ITT Internetall.** <mailto:Info@itt-sc.de>
- **Microsoft.** <http://www.microsoft.com/corpinfo/PRESS/Dec96/ntshw2pr.htm>
- **Opticom.** <http://www.opticom.de/>
- **Telos systems.** <http://www.zephyr.com/>
- **Worldspace.** <http://www.worldspace.com/>

- Traducciones de palabras que guardan interés con el contenido del artículo.

- Asignación de bit: bit allocation.
- Banco de filtros híbrido: hybrid filterbank.
- Banda crítica: critical band.
- Compatible hacia atrás: backwards compatible.
- Datos adicionales: ancillary data.
- Descompresión de trama: frame unpacking.
- Desplazamiento : shifting.
- Empaquetado de trama: frame packing.
- Escalado : scaling.
- Escaneo progresivo: progressive scan.
- Esquema - 1 : layer-1.
- Factor de escala: scalefactor.
- Gránulo : granule.
- Información lateral: side information.
- Muestra : sample.
- Nivel de enmascaramiento: masking threshold.
- Polifase : polyphase.
- Sonido envolvente: surround sound.

Sonido multilingüe: multilingual sound.

Tasa de bits: bit rate.

Tests de concordancia: conformance testing.

Tests de conformidad: compliance testing.

Glosario.

- Definiciones y siglas.

- AENOR: Asociación Española de Normalización. Organismo nacional que gestiona en España las normas ISO, entre otras.
- Aliasing: fenómeno indeseable que se produce al muestrear una señal por debajo de la frecuencia de Nyquist.
- Asignación de bit adaptativa: asignación variable de bits a las subbandas en tiempo y frecuencia de manera que se cumpla un determinado modelo psicoacústico.
- Asignación de ruido adaptativa: asignación variable de ruido de codificación a las bandas en frecuencia de manera que se cumpla un determinado modelo psicoacústico.
- CD: Compact Disc (Disco Compacto). Sistema de almacenamiento digital, sólo lectura, de unos 640 Mbytes de capacidad.
- Codificación entrópica: método de codificación que sin pérdida de información reduce la redundancia explotando las características de los métodos de longitud variable. La codificación Huffman es entrópica.
- CRC: Cyclic Redundancy Code (Código de Redundancia Cíclica). Código que explota las características cíclicas para ofrecer protección frente a errores mediante el empleo de redundancia.
- DCT: Discrete Cosine Transformation (Transformación Discreta del Coseno).
- DVD: Digital Video Disk (Disco de Vídeo Digital). Sistema de almacenamiento similar al CD, con la salvedad de que aumenta la capacidad hasta más de 10 Gbytes.
- Factor de escala: factor por el que se escala un conjunto de valores antes de la cuantización.
- FFT: Fast Fourier Transformation (Transformación Rápida de Fourier).
- Filtro de análisis: filtro en el codificador que transforma una señal de audio PCM de banda ancha en un conjunto de muestras de subbanda submuestreadas.
- Filtro de síntesis: filtro en el decodificador que reconstruye una señal de audio PCM a partir de las muestras de subbanda.

- HDTV: High Definition TeleVision (Televisión de Alta Definición).
- IMDCT: Inverse Modified DCT (DCT Inversa Modificada)
- ISDN: Integrated Services Digital Network (Red Digital de Servicios Integrados).
- ISO: International Standards Organization (Organización Internacional de Estándares).
- Mapeado: conversión mediante filtros de una señal de audio del dominio del tiempo al de la frecuencia.
- MDCT: Modified DCT (DCT Modificada).
- MIDI: Musical Instrument Digital Interface (Interface Digital para Instrumentos Musicales). Protocolo que permite la comunicación de ordenadores con diferentes elementos musicales y de éstos entre sí.
- Modelo psicoacústico: modelo matemático del comportamiento enmascarador del sistema auditivo humano.
- MPEG: Motion Pictures Experts Group (Grupo de Expertos en Imágenes en Movimiento).
- Muestras de subbanda: representación filtrada y submuestreada de la señal de entrada.
- NMR: Noise to Mask Ratio (Ratio Ruido a Máscara). Medida de la potencia del ruido respecto de la del nivel de la máscara utilizada en codificación subbanda.
- Recuantización: decodificación de las muestras de subbanda codificadas para recuperar las muestras originales.
- Segmentación adaptativa: subdivisión de la representación digital de una señal de audio en segmentos de tiempo variables.
- SMR: Signal to Mask Ratio (Ratio Señal a Máscara). Medida de la potencia de la señal respecto de la del nivel de la máscara utilizada en codificación subbanda.
- SNR: Signal to Noise Ratio (Ratio Señal a Ruido). Medida de la potencia de la señal respecto de la del ruido. Se utiliza como parámetro de calidad.
- WWW: World Wide Web (Red de Alcance Mundial). Sistema de intercambio de información desarrollado por el CERN en Suiza para mejorar las comunicaciones vía Internet.

Bibliografía.

- *Signal processing of speech*. F.J. Owens. Ed. Macmillan New Electronics. Nociones básicas de digitalización y codificación.

- Instituto Fraunhofer. <http://www.iis.fhg.de/>
FAQ sobre MPEG Audio esquema-3. / *departs/amm/layer3/sw*
Software (Winplay3, l3enc, l3dec). / *departs/amm/layer3/winplay3*
Especificaciones MPEG-1 y MPEG-2. / *departs/amm/layer3/sw*
Aplicaciones y futuro del estándar MPEG Audio. / *departs/amm/layer3/sw*
- ISO. <http://www.iso.ch/>
Información sobre las normas ISO-11172.
- MPEG3 Site. <http://www.mpeg3.com/>
Información general sobre el esquema-3.
- Notes on audio compression.
<http://fas.sfu.ca/cs/undergrad/CourseMaterials/CMPT479/material/notes>
Información sobre codificación de audio. Codificación subbanda. MPEG.
Modelos psicoacústicos.
- Layer3 personal tests.
<http://www.geocities.com/SiliconValley/Park/1486/l3test.htm>
Tests sobre el rendimiento del esquema-3.
- Philips Web Site. <http://www-us.philips.com/sv/newtech/mpeg/tecnolgy.htm>
Información sobre tecnología comercial MPEG Audio.