

5 CAPA DE AUDIO

Aunque en este proyecto no se desarrolla el decodificador del audio MPEG-2 considero de interés introducir algunos conceptos. La parte de la norma que recoge estas ideas es la ISO/IEC 13818-3.

5.1 Conceptos básicos de señal de audio

5.1.1 Necesidad de compresión

Al igual que hicimos con la señal de vídeo, la señal de audio debe también pasar a un formato digital. La tasa de bits de una señal de audio digital que ha sido codificada mediante PCM está alrededor del megabit por segundo. Como vemos, es muy inferior a la tasa del video digital sin comprimir. En aplicaciones que empleen una compresión media de la señal de audio no sería necesaria la compresión de audio. Pero, como el factor de compresión de vídeo está aumentando, sería conveniente comprimir este tipo de información también.

5.1.2 Técnicas de compresión de audio

La compresión de la señal de audio se basa en dos hechos. En primer lugar, en las señales de audio habituales, no todas las frecuencias están presentes simultáneamente. En segundo lugar, debido al fenómeno del enmascaramiento, el oído humano no puede diferenciar todos los detalles de la señal de audio, las frecuencias con mayor nivel enmascaran la información del resto de bandas. Este segundo factor se basa en que los sonidos reales no tienen una energía espectral uniforme.

Podemos distinguir dos técnicas fundamentales. Ambas se utilizan en estándar MPEG-2.

Codificación de sub-bandas

Consiste en dividir el espectro de la señal en bandas para filtrarlas o transformarlas, e incluye menos datos cuando está describiendo bandas en las cuales el nivel es menor. Cuanto menos sensibles seamos a bandas en particular, menos datos necesitan ser enviados. El número de bandas en las que se divide el espectro se elige dependiendo de la técnica de compresión que se empleará a continuación.

En el caso de MPEG se utiliza la compresión basada en el enmascaramiento, por eso es preferible que las bandas sean más estrechas que las bandas críticas del oído, y se requiere un gran número de bandas.

Las siguientes figuras muestran los diagramas de bloques de un codificador y de un decodificador de subbandas. En la entrada el rango de frecuencias es dividido en subbandas con la ayuda de un banco de filtros. Los datos descompuestos de la subbanda se organizan en bloques de tamaño fijo, antes del proceso de reducción (hay algunos codificadores que pueden utilizar bloques que se hacen más largos a medida que disminuye la frecuencia de la subbanda, no tienen porque ser de tamaño fijo). La ganancia de codificación se obtiene cuando la forma de onda de cada banda se pasa por un recuantificador.

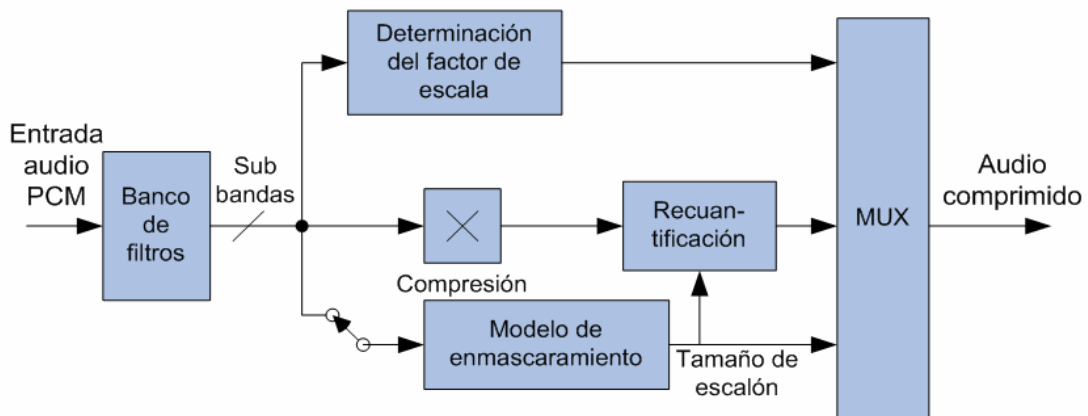


Figura 5.1 Esquema de un codificador de subbanda

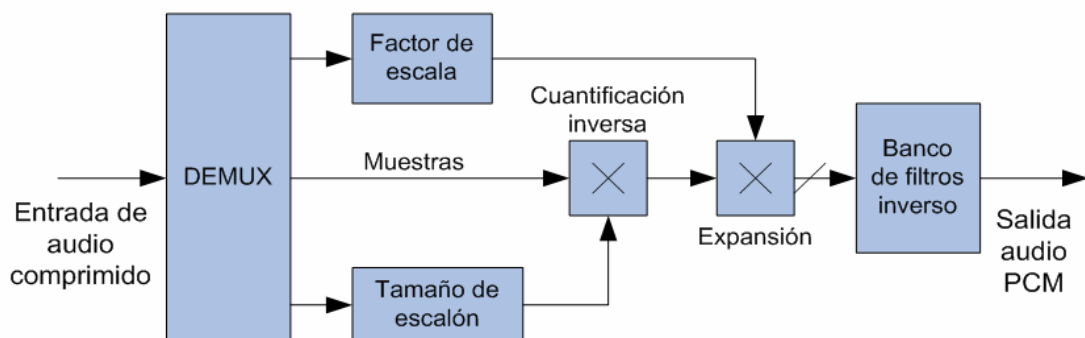


Figura 5.2 Esquema de un decodificador de subbanda

Codificación por transformación

Este tipo de análisis permite representar formas de onda a través de un conjunto de componentes armónicamente relacionados en amplitud y fase. La transformada de una forma de onda de audio varía de manera relativamente lenta.

En las transformadas prácticas se requieren bloques de muestras en lugar de cadenas interminables. Para conseguir trabajar correctamente se corta la forma de onda en pequeños segmentos solapados y, a continuación se transforman cada uno de ellos de forma individual.

5.2 Proceso de codificación en MPEG-2

El sistema de multiplexaje de MPEG-2, soporta cualquier número de canales de entrada de audio de forma que la velocidad de transporte seleccionada pueda soportar la suma de datos. Los usuarios tienen la flexibilidad para seleccionar su propio algoritmo de compresión de audio, tales como: Audio MPEG-2, MUSICAM, DOLBY AC-2 o AC-3. Los canales pueden ser configurados independientemente o en pares estéreo. La posibilidad de diferentes velocidades de audio es otra de las características del sistema, una vez más, la velocidad será asociada con la calidad.

MPEG-2 como principal novedad respecto a MPEG-1 incluye audio multicanal, en esta parte incluye dos estándares diferentes los cuales son completamente compatibles con MPEG-1, uno con compatibilidad hacia delante y el otro hacia atrás.

En las siguientes figuras podemos ver los diagramas de bloques del codificador y del decodificador empleados en MPEG-2.

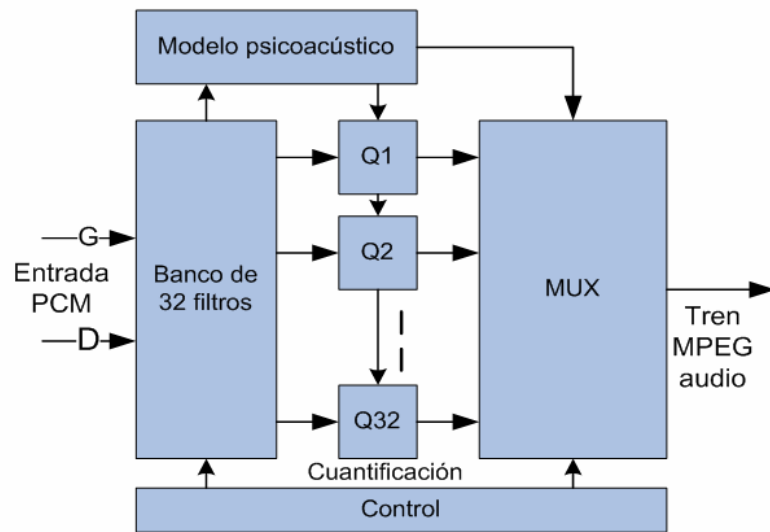


Figura 5.3 Esquema de un codificador de audio MPEG

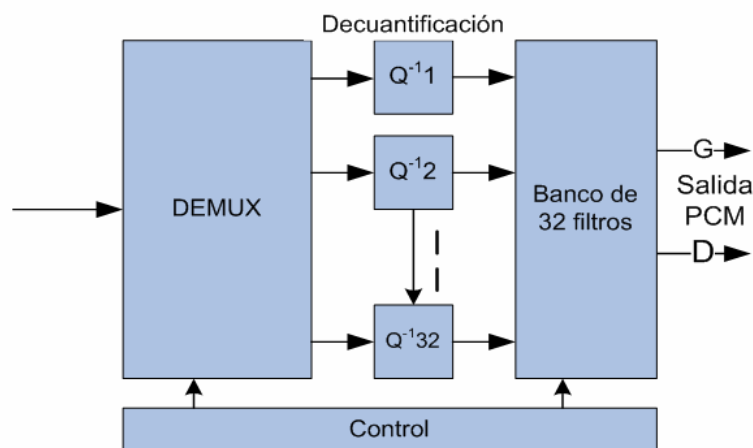


Figura 5.4 Esquema de un decodificador de audio MPEG

Las normas de MPEG definen tres capas de codificación:

- **Capa 1 (layer I).** Conocida como “pre-MUSICAM”, utiliza el algoritmo PASC de Philips.
- **Capa 2 (layer II).** Con su algoritmo “MUSICAM” permite obtener una calidad equivalente con un flujo menor.
- **Capa 3 (layer III).** Más conocida como “MP3” con un modelo psicoacústico, codificación de Huffman y análisis de la señal basado en DCT.

Desde el punto de vista de hardware y software, la más alta de las capas (III) incorpora el principal bloque para la construcción de las capas más bajas. Un decodificador de audio MPEG es capaz de decodificar tramas a partir de las tres capas.

5.2.1 Capa 1

Esta es la primera capa y la que menos prestaciones contiene. Por otra banda ofrece una complejidad menor y mayor velocidad dentro de la codificación/descodificación. Logra una tasa de 384kbps. Utiliza un banco de filtros polifásico basado en la Transformada Discreta del Coseno (DCT) dividiéndose en 32 subbandas igualmente espaciadas. Cada subbanda contiene 12 muestras.

5.2.2 Capa 2

Esta segunda capa mejora las prestaciones de la primera capa, con un aumento en la complejidad de cálculo. Logra una tasa de 192kbps, mejorando la tasa de transferencia de la primera.

Utiliza también un banco de filtros polifásico basado en la DCT y se divide en 32 subbandas igualmente espaciadas. Sin embargo usa tres conjuntos de muestras en el filtro (anterior, actual, próximo) consiguiendo 36 muestras por subbanda.

5.2.3 Capa 3

La tercera capa de codificación ofrece mejores prestaciones que las anteriores pero con una mayor complejidad en el cálculo y menos velocidad. Logra una tasa de 128kbps, mejorando la tasa de transferencia de las anteriores. Usa un banco de filtros híbrido conmutado, formado por un filtro polifásico DCT y por una transformación MDCT, permitiendo dividir el audio en bandas que se ajustan a las bandas críticas del oído. No hay igual espaciamiento entre subbandas. Integra un modelo psicoacústico que incluye los efectos del enmascaramiento tanto en la frecuencia como en el tiempo. Utiliza sofisticados esquemas de codificación por entropía y cuantificación no uniforme, incluidos la redundancia estéreo y los códigos Huffman creando tramas de longitud variable. Una de las aplicaciones más conocidas de esta capa es el formato de compresión de audio digital MP3.