

# Revisión del Estado del Arte de los Estándares de Codificación y Compresión de Audio MPEG y sus Aplicaciones

Herrera L.\*; Tipantuña C.\*\*

\*Facultad de Ingeniería Eléctrica y Electrónica, Escuela Politécnica Nacional, Quito, Ecuador

e-mail: [luis.alberto.herrera.lara@gmail.com](mailto:luis.alberto.herrera.lara@gmail.com)

\*\* DETRI – Facultad de Ingeniería Eléctrica y Electrónica, Escuela Politécnica Nacional, Quito, Ecuador

e-mail: [christian.tipantuna@epn.edu.ec](mailto:christian.tipantuna@epn.edu.ec)

---

**Resumen:** El presente artículo aborda una revisión del estado del arte de los estándares de codificación y compresión de audio de la familia *Moving Picture Experts Group* (MPEG). Desde los inicios de la era digital, la codificación de señales analógicas de audio en señales digitales y su posterior compresión han sido parte esencial de la actual era de la información. En este contexto, los estándares de codificación y compresión de audio MPEG han logrado una enorme acogida en aplicaciones multimedia, electrónica de consumo, producción de audio, telecomunicaciones, etc. El presente trabajo se enfoca en analizar los fundamentos teóricos de estos estándares así como también sus aplicaciones y los últimos avances en la mejora de los algoritmos usados en estos estándares y sus líneas abiertas de investigación.

**Palabras claves:** MPEG, codificación, compresión, señales analógicas de audio, codificación perpetua

**Abstract:** This article address a review of the state of art of standards of the *Moving Picture Experts Group* (MPEG) Family related to the codification and compression of audio signals. Since the dawn of the digital age, encoding analog audio signals into digital signals and its subsequent compression have been an essential part of the current era of information. In this context, the MPEG audio codification and compression standards have succeeded in multimedia applications, electronic consumer applications, audio production, telecommunications, etc. This paper focuses on analyzing the theoretical foundations, well as the applications and the latest developments in improving the algorithms used in these standards and their open research lines.

**Keywords:** MPEG, codification, compression, analog audio signals, codificación perpetua

---

## 1. INTRODUCCIÓN

Las técnicas de codificación de audio han sido y siguen siendo parte esencial de la actual sociedad de la información. Sin estas herramientas sería imposible imaginar los actuales avances aplicaciones multimedia, difusión de video, televisión digital, telecomunicaciones, entre otras. El objetivo principal de estos algoritmos es convertir señales analógicas de audio a formato digital garantizando la fidelidad de estas señales digitales con respecto a su original analógica. En este sentido, el desarrollo y mejoramiento de técnicas de codificación de audio así como también nuevos algoritmos y metodologías han suscitado el interés de muchos centros de investigación y universidades en donde expertos en codificación de audio trabajan en optimizar estas herramientas a fin

de conseguir resultados de alta fidelidad en la codificación de señales de audio analógicas en información digital [1–4].

El proceso de codificación de señales de audio analógicas en señales digitales consta básicamente de dos etapas fundamentales: el muestro y la cuantificación. En la etapa de muestreo se toma un conjunto de valores finito en el tiempo, a lo cual se conoce también como discretización de señales continuas. En esta etapa es necesario aplicar el teorema de *Nyquist* [4–6], el cual exige que para el proceso de discretización, que la frecuencia de toma de estos valores debe ser el doble o superior a la frecuencia de la señal discretizada para evitar los efectos de aliasing. En la etapa de cuantificación se representan los valores de la etapa de discretización con un cierto número de bits por muestra, donde reside el objetivo

---

Artículo recibido el XX, 2014; revisado XX Julio de 2014. (Escriba la fecha en que presento su documento para su revisión).

central de las técnicas de codificación de audio [5].

La etapa de cuantización como tal controla la fidelidad y tamaño del audio codificado en formato digital. En los inicios de la codificación de audio en señales digitales el método más popular a usar fue la Modulación por impulsos codificados (*Pulse Code Modulation - PCM*), donde la señal muestreada era cuantizada de manera uniforme. Este esquema dio origen a los primeros dispositivos de telefonía digital con señales discretizadas a  $8[kHz]$  y la música disponible en discos compactos con señales discretizadas a  $44.1[kHz]$ . En el caso de la música de audio con codificación para señales de audio estereofónicas se genera una tasa total de  $1.4[Mbits]$ , donde se puede apreciar ya, que para transmitir esta información en este formato se necesitaría un ancho de banda enorme. Este es el motivo central que promueve los esfuerzos de investigación en la optimización e innovación de algoritmos para la codificación de audio digital, la optimización del ancho de banda para la transmisión de información, en este caso, esto se traduce en reducir la información digital de las señales de audio codificadas a las mínimas necesarias garantizando la fidelidad de estas con respecto a sus originales analógicas [6, 7]. Gracias a esto es que hoy en día se disponen de servicios multimedios de transmisión de audio y video de gran fidelidad en redes de computadoras, transmisiones inalámbricas e internet [3-6].

En este trabajo se presenta una revisión introductoria a los Estándares de Codificación y Compresión de Audio *Moving Picture Experts Group* (MPEG) y sus Aplicaciones desde un punto de vista teórico - aplicativo. En la siguiente sección se analizan los principios sobre los cuales se definen los algoritmos usados en los estándares MPEG, luego en la siguiente sección se analizan las principales características y aplicaciones de los estándares MPEG-1, MPEG-2 y MPEG-4. También se presentan brevemente los algoritmos MPEG adicionales a los analizados. Luego se enumeran las principales líneas de investigación abiertas en el campo de la codificación de audio. Finalmente se anexan las conclusiones derivadas de la realización de este trabajo.

## 2. PRINCIPIOS BÁSICOS DE CODIFICACIÓN Y COMPRESIÓN DE AUDIO

La codificación de audio en la cual se basan los estándares MPEG se fundamenta en la filosofía de optimización de las señales digitales a través de la eliminación selectiva de las secciones de la señal de audio analógicas, las cuales son inaudibles para el oído humano. A estos criterios se los denomina también como técnicas de

codificación perceptual, ya que codifican solamente lo que es perceptible para el oído humano. La codificación perceptual se basa en los siguientes principios, el modelo psicoacústico (descripción físico - matemática) del oído humano, bandas críticas, el enmascaramiento temporal, el enmascaramiento frecuencial y finalmente la entropía perceptual. Estos conceptos son abordados a continuación en las siguientes subsecciones.

### 2.1 Umbral de Percepción del Oído Humano

El modelo psicoacústico de la percepción del oído humano fue definido por Fletcher en 1940. Este modelo muestra la capacidad de percepción del oído humano a sonidos de distinta frecuencia e intensidad o presión sonora en ambientes silenciosos [1, 2].

La Ecuación (1) describe la percepción del oído humano a valores de frecuencias  $f$ ,

$$T_a(f) = 3.64 \left( \frac{f}{1000} \right)^{-0.8} - 6.5e^{-0.6 \left( \left( \frac{f}{1000} \right) - 3.3 \right)^2} \dots \dots + 10^{-3} \left( \frac{f}{1000} \right)^4 \text{ [dB SPL]} \quad (1)$$

donde  $T_a(f)$  es el nivel de energía máximo permitido para las distorsiones producidas en la codificación de las señales de audio. Es decir, en cada frecuencia niveles mayores de intensidad de sonido a los descritos por la ecuación (1) dificultarán una correcta percepción de las señales de audio analógicas codificadas limitando así la fidelidad de las señales digitales con respecto a estas. En la práctica no se dispone de la información relativa a la intensidad de audio de las señales a codificar por lo que se asume que el nivel más bajo de la señal analógica a digitalizar es aproximadamente en  $4[kHz]$ , al cual se le asigna un nivel de amplitud de  $\pm 1[bits]$  de amplitud de voltaje, como nivel referencial para los reproductores de audio y sistemas de control de volumen [1, 2].

### 2.2 Bandas críticas

Las bandas críticas se definen como los rangos de frecuencias en los cuales la percepción del oído cambia abruptamente. Este fenómeno es característico en el funcionamiento de los sistemas sensoriales del oído y ocurre cuando se somete al oído a un primer tono el cual estimula un área determinada de sensores de audición, luego un tono adicional logra estimular un área mayor de sensores de audición, la cual incluye a la primera área estimulada. Dicho proceso de sobrelapamiento en la estimulación nerviosa opaca e incluso bloquea al

primer tono volviéndole inaudible, a pesar de que se eleve la intensidad de este con respecto al segundo tono, imposibilitando que el cerebro pueda distinguir dos tonos de frecuencia diferentes en la misma banda crítica.

Este fenómeno muestra que para el cerebro es más fácil distinguir tonos en baja frecuencia antes que en alta frecuencia. Las bandas críticas se estrechan en bajas frecuencias y agrandan en altas frecuencias debido a la fisiología del oído [1, 2]. Este proceso también puede ser interpretado como un banco de filtros muy solapados, en el cual las respuestas en cada banda crítica son asimétricas en amplitud y dependientes de la intensidad de la señal sonora. El ancho de banda de cada banda crítica es además no uniforme y se incrementa de forma no lineal con la frecuencia.

Para estimar el ancho de banda ( $BW_c(f)$ ) de cada banda crítica se usa la Ecuación (2),

$$BW_c(f) = 25 + 75 \left[ 1 + 1,4 \left( \frac{f}{1000} \right)^2 \right]^{0.69} \quad (2)$$

donde  $BW_c(f)$  se mantiene constante para rangos de  $100[Hz] - 500[Hz]$  aproximadamente y luego se incrementa alrededor de 20% en cada frecuencia central para valores superiores [8].

### 2.3 Enmascaramiento en temporal y frecuencial

Los fenómenos de enmascaramiento se producen cuando el oído está expuesto a dos o más sonidos simultáneos, donde uno de estos sonidos enmascara a los otros volviéndolos inaudibles. En la práctica, el nivel de intensidad de los sonidos del escenario descrito decidirá que sonido es audible y que sonido es perceptiblemente inaudible debido a este fenómeno de enmascaramiento, los sonidos más débiles enmascarados por los más fuertes. Existen dos tipos de enmascaramiento, el enmascaramiento temporal y el enmascaramiento frecuencial [1-3].

El enmascaramiento temporal se produce cuando un tono débil se encuentra cercano en el dominio del tiempo a un tono fuerte. La posición del tono débil define a su vez dos escenarios distintos, en los cuales este fenómeno se desarrolla de forma similar, el post-enmascaramiento y el pre-enmascaramiento. En el post-enmascaramiento si el tono débil llega luego del fuerte, el tono fuerte enmascarará al tono débil. En el pre-enmascaramiento en cambio, si llega primero el tono débil antes que el fuerte, el tono débil quedará igualmente enmascarado por el fuerte [2, 3].

En el fenómeno del enmascaramiento frecuencial en cambio pueden producirse dos casos en función de la intensidad sonora. En el primero los sonidos de baja frecuencia enmascaran a los sonidos de alta frecuencia y a su vez el caso contrario donde los sonidos de alta frecuencia enmascaran a los sonidos de baja frecuencia. Normalmente se produce el primer caso, ya que los tonos de alta frecuencia necesitaran una intensidad sonora considerablemente alta para poder enmascarar sonidos de baja frecuencia [2, 3]. En base a los fenómenos de enmascaramiento temporal y frecuencia, los algoritmos de codificación de audio deciden que parte de la señal analógica digitalizada codificar y que parte de esta señal descartar, ya que igualmente si se la codificará, esta sería inaudible para el oído humano [1-3].

### 2.4 Entropía Perceptual

La Entropía Perceptual (EP) combina el concepto clásico de entropía con los principios psicoacústicos básicos para evaluar los límites de comprensión de los segmentos de audio codificados. En otras palabras, evalúa hasta que límite es posible reducir la cantidad de información para que el cerebro pueda decodificar la información presente en el segmento de audio que se produzcan errores de comprensión [1, 2].

El proceso de estimación de la EP se realiza de la siguiente manera: la señal de audio analógica es primeramente inventanada y transformada al dominio de la frecuencia, en donde se aplican los enmascaramientos en base a las reglas del modelo psicoacústico. Luego se determinan el número de bits de resolución requeridos para la inyección de una señal sin ruido de procesamiento. La entropía perceptual es medida construyendo un histograma sobre muchas ventanas, escogiendo el peor caso para la actual medición [1, 2]. La unidad de medición de la EP es  $\left[ \frac{\text{bits}}{\text{muestra}} \right]$ , misma que viene expresada por la Ecuación (3).

$$PE = \frac{1}{N} \sum_{f=f_m}^{f=f_M} \max \left( 0, \log_2 \frac{S(f)}{\text{umbral}(f)} \right) \quad (3)$$

donde  $f_m$  es la frecuencia inferior mínima y  $f_M$  es la frecuencia superior máxima,  $N$  el número de componentes frecuenciales entre  $f_m$  y  $f_M$ ,  $S(f)$  es la densidad espectral de potencia de la señal y  $\text{umbral}(f)$  es el umbral estimado de enmascaramiento [8].

## 2.5 Tipos de codificadores

En base a los principios descritos en las sub-secciones anteriores existen las siguientes familias de codificadores: codificadores de transformación de componentes, codificadores en el dominio del tiempo, codificadores de bancos de filtros, codificadores de predicción lineal predictiva (LPC) y codificadores híbridos. Los codificadores de transformación se usan en sistemas de audio de alta fidelidad debido a su alta resolución de análisis espectral y temporal. Los codificadores de sub-banda en cambio optimizan la compresión de la señal en base a los principios psicoacústicos descritos anteriormente y a las redundancias estadísticas. Los codificadores adicionales funcionan mezclando las características de los codificadores de transformación y de sub-banda, a excepción de los basados en algoritmos de LPC, donde la codificación se hace modelando la señal de audio analógica digitalizada como una combinación lineal de muestras más una señal de error [1,2].

## 3. ESTÁNDARES MPEG

Los estándares MPEG de codificación y compresión de señales analógicas de audio basan su funcionamiento en los principios de codificación perceptual descritos anteriormente. La aplicación de de estos principios a los actuales algoritmos involucrados en los estándares MPEG se realiza de la siguiente manera. El modelo psicoacústico determina el máximo valor permisible de cuatización en bits para evitar introducir ruido en la señal codificada, así como también las secciones de la señal analógica a cuantificar que tienen características inaudibles al oído humano. Luego se realizan los enmascaramientos de los sonidos inaudibles en función de los parámetros determinados por el modelo psicoacústico. La siguiente etapa es normalizar la señal cuantizada y normalizarla, luego se identifican los tonos del modelo de enmascaramiento, para luego decimar y reorganizar estos marcadores. Finalmente se recalculan los coeficientes de los umbrales de enmascaramiento global [5-7, 9, 10]. La arquitectura básica de un codificador perceptual se define entonces por el análisis en el dominio tiempo - frecuencia, el modelo perceptual estimado, las etapas de cuantificación y codificación y el entramado de bits [8].

Los estándares MPEG poseen una arquitectura asimétrica en cuanto a la codificación - decodificación como tal, ya que la estructura del codificador presenta un grado de complejidad mayor que el decodificador. El objetivo de esto es claro, una vez codificadas las señales de audio analógicas, reproducirlas, es decir decodificarlas debe ser barato, lo cual vuelve a estos algoritmos altamente atractivos desde el punto de vista

comercial en aplicaciones relacionadas con la electrónica de consumo [6, 7, 9, 10]. Para la etapa de codificación, la señal de audio a aplicar al codificador MPEG debe estar en formato PCM en frecuencias de  $32[kHz]$ ,  $44.1[kHz]$  o  $48[kHz]$  con 16 bits de resolución por muestra. Debido a que es impráctico limitar la longitud de la señal a codificar, esta es dividida en tramas cuya duración se define como  $t_{trama} = \frac{1000}{f_m} \times m_{trama}$ , donde  $t_{trama}$  es el tiempo de duración de cada trama<sup>1</sup>,  $f_m$  la frecuencia de muestreo de la señal analógica y  $m_{trama}$ , el número de muestras de cada trama, lo cual dependerá de los algoritmos usados en la arquitectura del codificador [9, 10]. Un codificador MPEG funciona de la siguiente manera, la etapa de banco de filtros separa la señal de audio analógica en las bandas críticas, luego se evalúan los efectos psicoacústicos y se elimina la información irrelevante. La información relevante es luego codificada en la etapa de asignación de bits. Finalmente la etapa de codificación de trama formatea usando algoritmos de codificación como Huffman o Ziv-Lempel además de controlar el flujo de bits de la señal codificada, añadir información sobre el autor, códigos de acceso, marcas de agua para protección de derechos de autor, entre otras posibilidades [5-7].

El decodificador por su parte posee un arquitectura más sencilla en comparación al codificador. En el decodificador solamente se extrae la información formateada en el codificador, se extraen las bandas de la señal digital, se las reconstruyen por separado y finalmente se las unen para formar la señal de audio a reproducir. La codificación de audio como tal constituye solamente una parte de los estándares MPEG, dentro de la cual destacan los estándares MPEG-{1, 2, 3, 4}. En las siguientes subsecciones se describen los fundamentos básicos, características y aplicaciones de estos estándares.

### 3.1 MPEG-1

El estándar MPEG-1 también conocido como *ISO/IEC 11172*, es el primero de la familia de los estándares MPEG, el cual codifica imágenes en movimiento y audio (video) en velocidades generalmente de hasta  $1.5[Mbits]$ , aunque el estándar permite alcanzar velocidades más altas. Este estándar está estructurado en tres modos de funcionamiento y de complejidad creciente.

En *MPEG-1 Layer - 1* y *MPEG-1 Layer - 2* con velocidades de codificación de  $192[kBs]$  y  $128[kBs]$  para canales monoaurales y a  $384[kBs]$  y  $192[kBs]$  para canales estéreo [8]. La capa 1 y 2 presentan las siguientes

<sup>1</sup>En aplicaciones prácticas estos valores se encuentran en el rango de  $2[ms]$  a  $50[ms]$ .

características comunes: en la entrada de los datos, para agruparlos, se usa segmentación fija, luego de la segmentación se usan bancos de filtros polifásicos para dividir la señal analógica de audio en 32 sub-bandas. La asignación de bits se hace a partir del modelo psicoacústico del codificador. Finalmente la cuantización se hace en compasión de bloques y se añade un codificador de trama. La capa 2 realiza operaciones más optimizadas que la capa 1 debido a los valores de asignaciones de bits, para la capa 1 de  $\{0, 2, \dots, 15\}$  [bits] y en la capa 2 de  $\{3, 5, \dots, 7, 9\}$  [niveles] de cuantización, lo que reduce la velocidad de codificación optimizando la transmisión de audio codificado usando esta capa [8].

Este estándar incluye al formato *MP3*, denominado formalmente como *MPEG-1 Layer - 3* [6, 7, 9, 10]. *MP3* codifica a velocidades de  $64[kBs]$  en canales monoaural y a  $128[kBs]$  en canales estéreo. La capa 3 del estándar *MPEG-1* combina características de la capa 2 con codificación *Huffman* estática. A diferencia de las capas anteriores, en la capa 3 se usan filtros híbridos conmutados con resoluciones de  $\{576, 216, 192\}$  [lineas], cuantización no uniforme y tratamiento del ruido usando análisis por síntesis [8].

Sus principales aplicaciones de este estándar están orientadas a codificar imágenes en movimiento y audio en discos compactos en formatos de *Video CD*, *SVCD* y en baja calidad para *Video DVD*. También fue usado en aplicaciones relacionadas con televisión digital y televisión satelital.

### 3.2 MPEG-2

*MPEG-2* también conocido como *ISO/IEC 13818*, se define como un estándar mucho más amplio y de mayor atractivo para la codificación de video a velocidades de  $4 - 10[Mbits]$ . Ese estándar fue y es usado aun en la actualidad en la emisión de televisión digital de consumo masivo, cuya calidad es comparable con los sistemas analógicos *Phase Alternating Line - PAL*, *Sequential Color with Memory - SECAM*, *Over-the-Air Televisión Digital - ATSC*, *Digital Video Broadcasting - DVB*, *Integrated Services Digital Broadcasting - ISDB*, y *National Television System Committee - NTSC* [5-7, 9]. En comparación al estándar *MPEG-1*, este estándar presenta dos mejoras importantes: la codificación multicanal compatible hacia atrás *MPEG-2 BC* y el estándar derivado *MPEG-2 Advanced Audio Coder MPEG-2 - AAC*. *MPEG-2 BC* cambia los canales *L* y *R* (derecho e izquierdo) del estándar *MPEG-1* por los canales  $L_C$  y  $R_C$ , los cuales vienen definidos por las expresiones 4 y 5,

$$L_C = \frac{1}{1 + \frac{1}{\sqrt{2}} + a} \left[ L + \frac{C}{\sqrt{2}} + a \times L_S \right] \quad (4)$$

$$R_C = \frac{1}{1 + \frac{1}{\sqrt{2}} + a} \left[ R + \frac{C}{\sqrt{2}} + a \times R_S \right] \quad (5)$$

donde  $L_C$  y  $R_C$  son codificados finalmente usando el mismo estándar *MPEG-1*. El factor  $a$  puede tomar valores de  $\left\{ \frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{2\sqrt{2}}, 0 \right\}$ , por lo que este estándar puede reproducir audio digital de buena fidelidad a partir de una señal multicanal. En *MPEG-2 BC* se mantienen los mismos formatos de trama que en *MPEG-1*, los canales  $\{C, L_S, R_S\}$  se transmiten en campos auxiliares de la trama *MPEG-1* [1, 2, 8].

Para el caso de *MPEG-2 AAC* se elimina la compatibilidad hacia atrás y se introducen nuevos servicios estéreo multicanal. Este estándar presenta características adaptables al hardware de reproducción variando los recursos del canal y la calidad de reproducción deseada dependiendo del perfil de hardware disponible. El estándar recomienda la combinación de las siguientes herramientas: banco de filtros usando la transformada discreta de coseno levantado, estimación temporal del ruido, predicción como técnica de reducción de redundancias, escalamiento, cuantificación, codificación sin pérdidas usando códigos *Huffman* y el modelo psicoacústico similar al modelo 2 del estándar *ISO/IEC 11172-3* [1, 2, 8].

### 3.3 MPEG-3

El estándar *MPEG-3* no tuvo tanto éxito ni acogida como las versiones  $\{1, 2, 4\}$ , ya que presentaba redundancias con el estándar *MPEG-2*. Originalmente presentaba características escalables y de compresión multi-resolución orientado a aplicaciones de compresión de señales de video para televisión de alta definición con tasas superiores a  $10[Mbits]$  [6, 7].

### 3.4 MPEG-4

*MPEG-4* también denominado formalmente como *ISO/IEC 14496*. Este estándar mejora la codificación de la información a través del uso de algoritmos de compresión de mayor complejidad logrando así mayores radios de compresión que el estándar *MPEG-2*. Mas que un estándar *MPEG-4* es un conjunto de herramientas para aplicaciones de audio, voz y aplicaciones de sintetización de texto a voz, las cuales permiten realizar manipulaciones del sonido y composición *3D*, escalabilidad y procesamiento de efectos especiales, entre otros.

Este estándar está orientado mayoritariamente a aplicaciones multimedia para ordenadores para producción, distribución y acceso al contenido en televisión digital,

aplicaciones gráficas interactivas, aplicaciones multimedia, difusión de televisión en redes IP, dispositivos portátiles de juego, plataformas móviles de comunicaciones y entretenimiento, difusión de audio y video en internet, videoconferencias y difusión de radio y televisión satelital. La gran innovación de este estándar es convertir al decodificador de la información en un procesador de renderizado usando la información formateada para la descripción de formas tridimensionales y de textura de superficie. Se mejoran también los mecanismos de protección de propiedad intelectual, aplicaciones multimedia interactivas en ambientes Java. Las velocidades de codificación de este estándar rondan los  $64 - -256[kBits]$  [6, 7, 9–11].

### 3.5 Estándares MPEG Adicionales

Adicionalmente a los estándares descritos en la sección anterior existen variaciones y extensiones de los mismos que los complementan en sus aplicaciones en la codificación de audio y video. En [2, 7, 9–11, 15]

- *MPEG-7*, estándar usado en la descripción y búsqueda de audio digital.
- *MPEG-21*, estándar usado en Aplicaciones Multimedia.
- *MPEG-A*, estándar usado para aplicaciones específicas en formatos especiales de audio y video.
- *MPEG-B*, agrupa un conjunto de sistemas de estándares usados en aplicaciones específicas.
- *MPEG-C*, agrupa un conjunto de sistemas de estándares usados en aplicaciones de video.
- *MPEG-D*, agrupa un conjunto de sistemas de estándares usados en aplicaciones de audio.
- *MPEG-DASH*, estándar usado para la difusión de video en internet.
- *MPEG-E*, un estándar (M3W) que provee el soporte para la descarga y ejecución de aplicaciones multimedia.
- *MPEG-H*, estándar usado en aplicaciones de codificación de video de alta eficiencia.
- *MPEG-M*, estándar usado en módulos de dispositivos móviles con tecnología (PCI) para el empaquetado y reúso de tecnologías MPEG.
- *MPEG-U*, estándar usado en el mejoramiento de la interfaz de usuario.
- *MPEG-V*, estándar usado en el intercambio de datos en realidad virtual.

## 4. LÍNEAS DE INVESTIGACIÓN ABIERTAS

A pesar que el estado de arte actual de los algoritmos de codificación de señales de audio es altamente superior al mostrado por estos en los orígenes de estas herramientas, existen aun problemas por solucionar y mejoras por desarrollar en cuanto a la optimización de la compresión de la información digital y el mejoramiento de la fidelidad de las señales de audio codificadas. En [1, 9] se analizan varias de las posibles líneas de investigación abiertas en el campo de la codificación y compresión de audio de entre las cuales destacan las siguientes:

- Nuevos Algoritmos de codificación usando bancos de filtros basados en *Wavelets*, *Choi-Williams*, *Zhao-Atlas-Marks*, *Born-Jordan* [12] y *Matching Pursuit* [13, 16–18].
- Esquemas alternativos a la codificación en subbandas, usando esquemas de codificación paramétrica.
- Predicción lineal con excitación por código(CELP) con codificación multimodo, usando códigos *Huffman*.
- Empaquetamiento de códigos usando *Wavelets adaptadas*.
- Codificación de tramas usando transformadas diferenciales(DPAC).

## 5. CONCLUSIONES

- Los algoritmos de los estándares MPEG basan su funcionamiento en los conceptos de codificación perceptual de audio, la cual es una técnica de codificación con pérdidas, donde la cantidad de información se optimiza descartando la información inaudible o imperceptible al oído humano de acuerdo a sus características temporales y frecuenciales (fenómenos de enmascaramiento).
- Los estándares MPEG como tal no poseen una definición estricta que los limite en cuanto a rendimiento, al contrario, son estándares abiertos donde solamente se realizan recomendaciones en cuanto al flujo de bits, para garantizar la compatibilidad entre reproductores de audio y recomendaciones en cuanto a la implementación de codificadores, esto ha permitido fomentar la investigación e innovación en nuevos algoritmos de codificación de audio manteniendo al mismo tiempo la compatibilidad entre codificación y decodificador independientemente de la estructura interna usada.

- Uno de los problemas más complejos en la definición de algoritmos y estándares de codificación es la compatibilidad entre algoritmos. Para esto, MPEG define sus estándares en capas, en donde cada capa define un codificador más sofisticado que el anterior. Los estándares de capa 1 solamente pueden reproducir codificaciones de capa 1, mientras que los estándares de capa 3 pueden reproducir codificaciones de capa 2 y capa 1.
- Los algoritmos de codificación de audio y estándares se encuentran actualmente en un grado avanzado de desarrollo, sin embargo existen aun temas por abordar en la optimización y mejoramiento de los procesos de codificación y fidelidad de las señales de audio.

#### REFERENCIAS

- [1] Painter, T.; Spanias, A., "A review of algorithms for perceptual coding of digital audio signals," Digital Signal Processing Proceedings, 1997. DSP 97., 1997 13th International Conference on , vol.1, no., pp.179,208 vol.1, 2-4 Jul 1997, doi: 10.1109/ICDSP.1997.628010.
- [2] J. Youngquist Robert, "Sampling Frequency Considerations," Journal of the Audio Engineering Society, pp. 54.
- [3] M. Willcocks, "A Review of Digital Audio Techniques," Journal of The AES, Jan-Feb, 1978, Volume 26, pages 56-64.
- [4] Heaslett A., "Some Criteria for the Selection for Sampling Rates in Digital Audio Systems," J. Aud. Eng. Soc., 26: 66, 68, 70, Jan./Feb. 1978.
- [5] Johnston, J.D., "Perceptual Audio Coding - A History and Timeline," Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on , vol., no., pp.2085,2087, 4-7 Nov. 2007, doi: 10.1109/ACSSC.2007.4487605.
- [6] Spanias, A., Painter, T., Atti, V., *Audio Signal Processing and Coding*, ISBN: 9780470041963, 2006, Wiley Press.
- [7] Brandenburg K., "Mp3 and AAC Explained," AES 17th International Conference on High Quality Audio Coding, 1999.
- [8] Candeas Vera P., *Desarrollo de Técnicas de Codificación de Audio Basadas en modelos de señal paramétricos*, Tesis Doctoral, Departamento de Teoría de la Señal y Comunicaciones, Universidad de Alcalá, 2006.
- [9] Rincón Rivera, David. *MP3 : sonido digital al alcance de todos*. "Buran", 1999, núm. 14, p. 4-12.
- [10] Noll, P., "MPEG digital audio coding," Signal Processing Magazine, IEEE , vol.14, no.5, pp.59,81, Sep 1997, doi: 10.1109/79.618009
- [11] Peter Noll, "MPEG Digital Audio Coding Standards, Chapter in: IEEE Press/CRC Press, The Digital Signal Processing Handbook" (ed.: V.K. Madisetti and D. B. Williams), pp. 40-1 - 40-28, 1998
- [12] Manas A., Neha M., *Audio Compression in MPEG Technology*. International Journal of Scientific and Research Publications, Volume 3, Issue 12, December 2013, ISSN 2250-3153.
- [13] Davis Pan. 1995. A Tutorial on MPEG/Audio Compression. IEEE MultiMedia 2, 2 (June 1995), 60-74. DOI=10.1109/93.388209 <http://dx.doi.org/10.1109/93.388209>.
- [14] Peter L. Stanchev. 2011. *Multimedia standards. history. state of art*. In Proceedings of the Third international conference on Future Generation Information Technology (FGIT'11), Tai-hoon Kim, Hojjat Adeli, Dominik Slezak, Frode Eika Sandnes, and Xiaofeng Song (Eds.). Springer-Verlag, Berlin, Heidelberg, 39-42. DOI=10.1007/978-3-642-27142-76 <http://dx.doi.org/10.1007/978-3-642-27142-76>.
- [15] Brandenburg, K.; Faller, C.; Herre, J.; Johnston, J.D.; Kleijn, W.B., "Perceptual Coding of High-Quality Digital Audio," Proceedings of the IEEE , vol.101, no.9, pp.1905,1919, Sept. 2013 doi: 10.1109/JPROC.2013.2263371.
- [16] Roberto Herrera Lara, Robin Álvarez Rueda, Alexis Andrango Pillajo, "Evaluación de las técnicas tiempo-frecuencia por medio de un equipo de adquisición de datos y un computador". Avances en Ciencias e Ingeniería, 2011, Vol. 3, No. 1, Pags. C33-C39. ISSN 1390-5384 .
- [17] S. G. Mallat and Z. Zhang, "Matching Pursuits with Time-Frequency Dictionaries", IEEE Transactions on Signal Processing, December 1993, pp. 3397-3415.
- [18] X. A. Calderón, C Montenegro, C Mullo, C Samaniego, "Desarrollo de una Aplicación Cliente/Servidor para un Wall View en base a la Plataforma Cruzada Opensource FFMPEG (Colección de Software Libre que puede Grabar, Convertir y hacer Streaming de Audio y Vídeo)", La Revista Politécnica, EPN Journal, Vol. 33, No 3., Enero 2014.