

MP3 Magia o escoria

En la disputa actual por el liderazgo en la alta resolución y fidelidad dentro del audio digital, como por ejemplo: SACD (Super Audio Compact Disc), 24 bits a 192 kHz, DSD (Direct Stream Digital), etc.; encontramos formatos que disienten con estos preceptos. La base de esta dicotomía está radicada en la necesidad de ahorrar bits por segundo para una transferencia determinada. Especialmente si hablamos de audio para aplicaciones como:

- ❑ DAB: Digital Audio Broadcasting (Transmisión de Audio Digital).
- ❑ Audio por Internet.
- ❑ Distribución y emisión de programas de radio.
- ❑ Radiodifusión por cable y satélite.
- ❑ Televisión por cable y satélite.
- ❑ Telefonía.
- ❑ Sonido en varios canales (DTS¹, SDDS², Dolby Digital).
- ❑ DCC: Digital Compact Cassette (Cassette Compacto Digital).
- ❑ DAW: Digital Audio Workstations (Estaciones de Trabajo Digital).
- ❑ CD-I: Compact Disk Interactive (Discos Compactos Interactivos).
- ❑ Comentarios, reportes deportivos y noticias por líneas ISDN³.
- ❑ Sistemas de anuncios públicos, grabación de audio en estado sólido (memo-recorders, juguetes, paging, etc.).
- ❑ MD: Mini Disk (Minidisco).
- ❑ HDTV: High Definition Television (Televisión de Alta Definición).

La vedette de estos formatos es el popular MP3 (MPEG-1 layer 3 o nivel 3, "MP3" es la extensión que utiliza Windows para identificar el tipo de archivo), sin embargo, existe una cantidad considerable de ellos y cada uno se creó con un fin específico.

Los principios para reducción de bits, también llamado "compresión" (no confundir con la mengua de rango dinámico), se encuentran en un trabajo que se originó en el año 1983 por una empresa Argentina que desde 1968 venía desarrollando y fabricando consolas y equipamiento para radiodifusión. Me refiero a la firma Solidyne, de la cual el Ing. Oscar Bonello es presidente y fundador.

En los inicios, lo que se trataba de lograr era la incorporación de sonido a la PC, que debido a las características tecnológicas, en lo que a hardware se refiere, resultaba una tarea imposible⁴.

¹ DTS: siglas de Digital Theatre System (Sistema Digital para Teatros). Se trata de un formato multipista para cine con 6 canales de audio.

² SDDS: siglas de Sony Dynamic Digital Sound (Sonido Digital Dinámico de Sony). Es un formato abierto que permite hasta 8 canales discretos de audio.

³ ISDN: siglas de Integrated Services Digital Network (Red Integrada de Servicios Digitales). Su aplicación más característica es la de enlace entre proveedores de Internet con la Red de Internet misma.

⁴ Realizando un sencillo cálculo notaremos que, un CD común tiene 44100 muestras por segundo de dos palabras de 16 bits, por lo tanto, para almacenar 1 minuto de audio necesitaremos aproximadamente 10 MB de capacidad. Tengan en cuenta que en esa época (1983) la capacidad de un disco rígido de una PC XT, era de 10 MB; y mejor no hablar de los costos (3000 dólares).

De ésta necesidad surgió la primer norma para reducción de bits del mundo, llamada ECAM: Enhanced Compressed Audio Modulation (Modulación de Audio Comprimido Mejorado), que estaba incorporada a las placas digitalizadoras "Audi-com" de la misma empresa.

En 1987, fecha en la que Solidyne estaba por lanzar la quinta generación de placas que fabricó, comenzaron los estudios sobre reducción de bits para transferencia de audio en el instituto alemán Fraunhofer, de cuya conclusión resultó el MPEG⁵. La evolución de los mismos hizo eclosión en noviembre de 1992 a instancias del experto italiano Leonardo Chiariglione con el primer estándar de compresión para audio y video aceptado por ISO: MPEG-1⁶.

Con la necesidad de obtener reducción de bits en formatos multicanal (5.1, 7.1, etc.) y para frecuencias de muestreo más bajas que las del MPEG-1, en noviembre de 1994 finalizó MPEG-2 BC⁷. La calidad del audio codificado por MPEG-2 BC no es buena, por eso, en abril de 1997 surgió MPEG-2 NBC⁸.

En noviembre de 1998 se terminó MPEG-4, que está preparado para trabajar con muy bajas tasas de transferencia, tiempos de retraso despreciables y manteniendo alta calidad de audio. El concepto de su estructura difiere relativamente poco de sus predecesores.

Otros formatos para compresión de audio en forma de datos son:

- Dolby AC-2 y AC-3: Audio Coding (Codificación de Audio).
- Sony ATRAC: Adaptive transform Acoustic Coding (Codificación por Transformación Acústica Adaptable).
- AT&T PAC/MPAC: Perceptual Audio Coder (Codificador de Audio Perceptible)/ Multichannel Perceptual Audio Coder (Codificador de Audio Perceptible Multicanal).
- APT-X100.

El principio de funcionamiento de la mayoría de estos estándares se basa en una secuencia de pasos o procesos que modifican la señal de entrada de la siguiente manera:

- 1) Un banco de filtros que divide al espectro en bandas y submuestras, obtiene de esta forma paquetes de señal seleccionados en el dominio de la frecuencia y el tiempo.

⁵ MPEG: siglas de Moving Pictures Expert Group (Grupo especialista de cine). Se estableció en 1988 para especificar esquemas de codificación de datos con baja tasa de transferencia en video y audio digital.

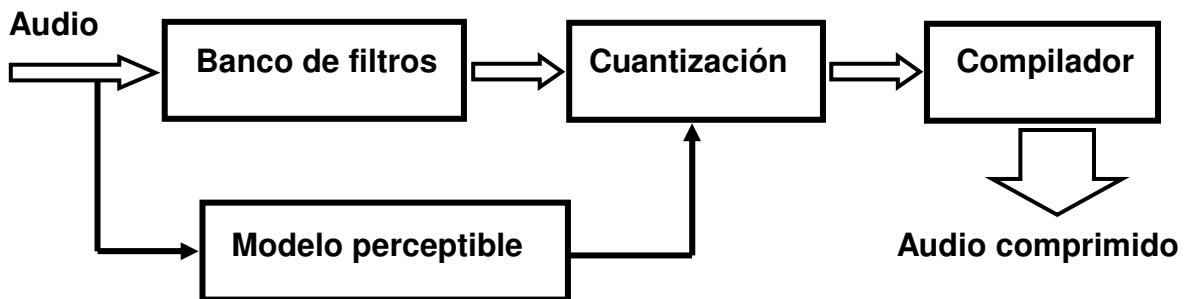
Para ser más precisos el cuerpo de estandarización original y completo es: ISO/IEC JTC1/SC29/WG11 que se conoce como MPEG. Donde ISO es International Standardization Organization (Organización Internacional de Estandarización). IEC: International Electrotechnical Commission (Comisión Internacional Electrotécnica). JTC1: Technical Committee J1 (Comité Técnico J1). SC: SubCommittee (Subcomité). WG: Work Group (Grupo de Trabajo).

⁶ El formato MPEG1 se conoce como: IS 11172-3. International Standard (Estándar Internacional).

⁷ El formato MPEG-2 BC se conoce como: IS 13818-3. BC: Backwards Compatible (Compatible con formatos anteriores).

⁸ El formato MPEG-2 NBC se conoce como: IS 13818-7. NBC: Non Backwards Compatible (No Compatible con formatos anteriores). Esta norma es la conocida MPEG-2 AAC. Advanced Audio Coding (Codificación de Audio Avanzada).

- 2) Usando la misma señal de entrada o la salida del banco de filtros, se calculan los umbrales de enmascaramiento en base a reglas psicoacústicas ya determinadas. Esta etapa se llama "Modelo perceptible".
- 3) Las componentes espectrales se cuantizan y codifican en base a los algoritmos que se obtienen del modelo perceptible.
- 4) Por último el compilador es el que reúne todos estos parámetros y genera una cadena de bits, que es la señal codificada o comprimida a utilizar.



Hablemos brevemente sobre las etapas mencionadas:

- 1) Banco de filtros: son los que proveen la base de análisis del sistema. Se utiliza la transformada discreta de Fourier⁹ y en algunos casos se acopla una pequeña etapa que puede procesar en el dominio del tiempo las muestras que son adyacentes. Matemáticamente hablando, todas las transformaciones discretas usadas hoy en día para audio se las puede llamar banco de filtros. Los filtros que se subdividen pueden estar contemplados como una transformada de 1 a n componentes.
- 2) Modelos perceptibles: son los encargados de estimar los umbrales de enmascaramiento *imprescindibles* en toda codificación del tipo perceptible. La precisión de esta etapa va a determinar la calidad subjetiva que se perciba a bajas tasas de transferencia. En el caso más simple se utiliza un modelo estático. Otro sistema calcula la relación señal ruido necesaria para cada banda derivada de las curvas de enmascaramiento.

Un modelo muy simple para asignación de bits se obtiene de la siguiente forma:

$$n \text{ bits} = [27 \text{ dB} \times (B_s - B_i) / 6,02 \text{ dB}]$$

Donde:

B_s es el límite superior de la banda y B_i es el límite inferior de la banda. Lo que la fórmula indica es, que el número de bits por banda se obtiene del enmascaramiento si la señal que enmascara es ruido de alta frecuencia del final de la banda. Otros modelos más avanzados estiman una relación señal-enmascaramiento dependiente del tiempo, por cada banda usada en el codificador.

- 3) Es la etapa que realiza el mayor trabajo de reducción de bits en un codificador. Actualmente existen dos métodos muy utilizados que son:

⁹ Con la transformada de Fourier $F\{s(t)\}$ se lleva a cabo, con ayuda de la integral directa de Fourier, un desarrollo de la función temporal $s(t)$ en un espectro continuo (densidad espectral) $S(\omega)$, en el cual la frecuencia corresponde a la densidad del espectro.

- : Un grupo de valores ordenados en el dominio de la frecuencia o el tiempo se normaliza al máximo valor absoluto (el valor de la normalización se llama factor de escala). Los valores de cada bloque se cuantizan con un tamaño que se obtiene de la asignación de bits de ese bloque, dado por el modelo perceptible.
 - *Asignación de ruido seguida por cuantización escalar y codificación Huffman:* En este método no se usa asignación de bits, en cambio, una cantidad de ruido permitido, igual al umbral de enmascaramiento estimado, se calcula para cada banda por medio del factor de escala. El factor de escala se usa para llevar a cabo una coloración del ruido de cuantización, o sea, modificar el tamaño del paso de cuantización para todos los valores, con el factor de escala de cada banda. Los valores ya cuantizados se codifican usando codificación Huffman¹⁰.
- 4) El compilador simplemente arma la cadena de bits que se van a transmitir. Hagamos un repaso de las familias de MPEG:

Formato	Frec. de muestreo	Canales	Complejidad
MPEG 1	32, 44.1, 48 KHz	mono / estéreo	Layer 1 (nivel 1) Layer 2 (nivel 2) Layer 3 (nivel 3)
MPEG 2 BC	16, 22.05, 24 KHz	mono / estéreo	Layer 1 (nivel 1) Layer 2 (nivel 2) Layer 3 (nivel 3)
	32, 44.1, 48 KHz	Multicanal	Layer 1 (nivel 1) Layer 2 (nivel 2) Layer 3 (nivel 3)
MPEG 2 NBC (AAC)	32, 44.1, 48 KHz	mono, estéreo y Multicanal	Layer 1 (nivel 1) Layer 2 (nivel 2) Layer 3 (nivel 3)

Recordemos que el tan difundido y conocido por todos "MP3" es el MPEG 1 layer 3. Veamos algunas diferencias cualitativas que se presentan en esta familia:

Formato	Tasa trans.	Canales	Calidad subjetiva	Ancho de banda
MPEG 1 Layer 3	128 Kbits/s	Estéreo	Calidad muy cercana al CD	≥ 16 KHz
	96 Kbits/s	Estéreo	Calidad cercana al CD	15 KHz
MPEG 2 Layer 3	64 Kbits/s	Estéreo	Calidad cercana a FM	11 KHz
	32 Kbits/s	Mono	Mejor que AM	7.5 KHz
	16 Kbits/s	Mono	Mejor que onda corta	4.5 KHz

¹⁰ Método de codificación que posee un promedio de reducción de 1,5 bits por muestra y por canal. Básicamente usa los bits necesarios para el valor, más uno de signo, eliminando la redundancia de ceros. En la práctica la forma de onda del audio representada como palabras digitales, tienden a seguir una amplitud de probabilidad estadística conocida como Laplaciana; esta propiedad se utiliza en la codificación Huffman para la predicción de los valores de cada bloque.

Ahora cotejemos algunas diferencias entre otros formatos.

Tabla 3: Comparación entre formatos				
Formato	Tasa trans.	Filtro	Resolución en frecuencia	Uso principal
MPEG 1 L1	32~448 Kbits/s	PQMF	750 Hz	DCC
MPEG 1 L2	32~384 Kbits/s	PQMF	750 Hz	DAB, CD-I, DVD
MPEG 1 L3	32~320 Kbits/s	PQMF / MDCT	41.66 Hz	ISDN, Audio en internet
ECAM	20:1	4 bandas.		DAB
Dolby AC2	128~192 Kbits/s /ch	MDCT / MDST	93.75 Hz	Conexiones punto a punto
Dolby AC3	32~640 Kbits/s	MDCT	93.75 Hz	HDTV, Cable, DVD
ATP-X100	1:4 fija	QMF		Estudios de grabación
Sony ATRAC	140 Kbits/s	QMF / MDCT	46.87 Hz	MD

Lo que representa el rótulo de "complejidad", última columna de la tabla 1, es el grado de dificultad que tienen los algoritmos del banco de filtros y los de cuantización, para llevar a cabo su proceso. Según el estándar ISO, el layer 1 y 2 usa un filtro PQMF dividiendo la señal entrante en 32 bandas que después se cuantizan controladas por un modelo psicoacústico para luego codificarse con el sistema Bloque de punto flotante. La aplicación más usual del grado de complejidad layer 1 (nivel 1), son aquellas prácticas donde no se necesiten tasas de transferencia muy bajas, por ejemplo, audio hogareño, DCC, discos magneto ópticos, etc. En el layer 2 (nivel 2) hay más merma de datos eliminando la redundancia de los factores de escala y además, la cuantización es más precisa. El layer 3 (nivel 3) consiste en módulos de ASPEC (Adaptive Spectral Entropy Coding - Codificación por Entropía Espectral adaptable), que le brindan alta eficiencia en la codificación. El banco de filtros se denomina híbrido, pues utiliza las 32 bandas del PQMF que a su salida divide cada una de éstas en 18 sub bandas mediante MDCT, logrando así mayor resolución en frecuencia. Layer 3 emplea una información adicional llamada "Bit de depósito" que ayuda al sistema (codificador y decodificador) a resolver problemas cuando hay demanda en la tasa de transferencia y es el único que puede trabajar con tasa variable sin violar el estándar ISO.

La nomenclatura de los filtros utilizados en estas normas son:

QMF: Quadrature Mirror Filter. (Filtro Espejo en Cuadratura).

PQMF: Polyphase Quadrature Mirror Filter. (Filtro Espejo en Cuadratura Polifásico).

MDCT: Modified Discrete Cosine Transforms. (Transformada Discreta de Coseno Modificada).

MDST: Modified Discrete Sine Transforms. (Transformada Discreta de Seno Modificada).

Un suplemento que reduce aún más la cantidad de datos, es el concepto de Joint Stereo Coding (Codificación Estéreo Combinada). El principio de su funcionamiento se basa en disminuir la cantidad de información espacial (imagen estéreo), sin introducir distorsión audible. Hay dos factores que permiten esta merma: la redundancia y los agregados de datos que acrecientan la percepción estéreo. Estos últimos se conocen como "stereo irrelevancy" (impertinencia estéreo), y se manifiestan entregando vestigios para obtener información espacial, a partir de sonidos de alta frecuencia tomados desde los fragmentos de máxima energía.

Existen dos métodos para el sistema estéreo combinado:

- Codificación por matriz estéreo (M/S): La técnica utilizada es similar al matrizado para transmisiones de FM estéreo. Así, en vez de transferir los dos canales, se trasladan la suma (M) y por otro lado la diferencia (S) de ambos canales.
- Codificado estéreo por intensidad: Para cada banda que se transmite usando el modo de intensidad estéreo, sólo la penetración de la información se captura. Estos datos de direccionalidad se transmiten por medio del codificador con un factor de escala independiente para cada canal. De esta manera, se envía sólo la envolvente de la energía de ambos canales.

Haciendo un pequeño resumen de lo que hasta aquí he mencionado, se puede apreciar que del material original queda sólo aquello que, en teoría, el oído tiene la habilidad de discernir para entender la información que se desea transmitir. Dicho de otra forma, las componentes espectrales que son enmascaradas¹¹ por las de mayor amplitud se eliminan. De la misma manera son anuladas las componentes que en el dominio del tiempo están próximas a una de mayor amplitud por enmascaramiento temporal¹². Toda esta información procesada se embla mediante algún sistema de codificación que elimine la redundancia de bits (por ejemplo: Huffman) y luego se envía a su destino.

Los conceptos de enmascaramiento, que son la piedra fundamental en la que se basan los algoritmos para la reducción de datos de audio, se estudian en una disciplina llamada psicoacústica y se los conoce desde 1924.

El decodificador al recibir el paquete, o sea, la cadena de bits codificados, ejecuta la conversión o decodificación para poder reproducir el audio. Cabe acotar que

¹¹ Cuando una señal senoidal pura estimula la membrana basilar del oído interno, abarca una región que comprende más de una celda sensorial, por lo tanto se halla excitando la zona en varias componentes del espectro. Dependiendo de la amplitud y la frecuencia de dicho estímulo, va a existir un rango de frecuencias dentro del cual las componentes cercanas (en frecuencia) van a ser enmascaradas por la de mayor amplitud. En consecuencia, el oído no tiene la habilidad de "escuchar" todas aquellas componentes que se encuentren por debajo del umbral que genera el tono enmascarante.

¹² El proceso está basado en un fenómeno psicoacústico del oído que funciona de la siguiente manera: cuando se recibe una excitación de un valor de presión sonora elevado, el oído interno incrementa el nivel del umbral de audición. Toda señal posterior con diferencia de pocos milisegundos a la excitadora, y que se encuentre por debajo del umbral de audición modificado, simplemente no será percibida. Este principio depende de varios factores: el nivel de excitación, la frecuencia, su ancho de banda, el tiempo de estimulación, la diferencia temporal entre señales, etc.

aunque se lo convierta nuevamente al formato PCM lineal¹³ mediante programas específicos para tales fines, la información inicial ya *no existe*; esto es, que *no hay forma* de obtener otra vez el material original. Por lo tanto se realiza compresión destructiva.

Existen métodos "no destructivos" para comprimir señales o datos, que no alteran el contenido primitivo del audio (una vez descomprimido), pero la reducción de bits que se logra es poco alentadora (en el caso del audio). Uno de los formatos más populares es el "ZIP"¹⁴ y a modo de ejemplo muestro algunos resultados en base a ésta norma:

Tipo de información del audio (44.1KHz - 16bits)	Tamaño original	Tamaño comprimido	Relación de compresión
Tono puro senoidal 1KHz (10 s)	862 KB	8KB	99%
Redoble de batería (2,4 s)	210 KB	168 KB	23%
Richard Strauss "Así habla Zaratrustra" (fragmento 1:48)	19,36 MB	16,34 MB	16%

Obviamente la cantidad de compresión que se alcance estará de acuerdo con el contenido que posea la señal a compactarse; en consecuencia, mientras más estocástico sea el material, menor será el grado de reducción. Por otra parte, en el caso del MPEG-1 la cantidad de reducción de bits llega a valores grandes para un material como el audio, y no es tan benéfico para señales periódicas (tono puro).

Tipo de información del audio (44.1KHz - 16bits)	Tamaño original	Tamaño comprimido	Relación de compresión
Tono puro senoidal 1KHz (10 s)	862 KB	157 KB	82%
Redoble de batería (2,4 s)	210 KB	38 KB	82%
Richard Strauss "Así habla Zaratrustra" (fragmento 1:48)	19,36 MB	1,75 MB	91%

¹³ Pulse Code Modulation (Codificación por modulación de pulsos). Norma de audio sin compresión de datos, es la señal digitalizada intacta como se imprime en un CD (en el caso que tenga frecuencia de muestreo de 44,1 KHz y 16 bits de resolución). La extensión conocida en los programas de computación es .WAV.

¹⁴ El método es el PKZIP (muy familiar en los ámbitos de la computación) y se basa en un proceso que elimina la redundancia de información. Por ejemplo si tomamos una muestra de 4KHz con resolución de 20 bits, muestreada a 48KHz; las primeras 5 palabras digitales serán:

1. 00000000010000110000
2. 00000000011000010000
3. 00000000011001100000
4. 00000000010011110000
5. 0000000001000110000

En este caso se pueden transmitir los 11 bits que van cambiando acompañados por la información que indique la precedencia de 9 bits con valor "0".

Otros formatos similares al zip son: z; gz; tz; taz; tgz; lzh; arj; arc; tar; uu; uue; xxe; bhx; b64; hqx.

La calidad subjetiva que se logra con la norma MPEG-1 layer 3 (tabla 2) a una tasa de transferencia relativamente baja (128 Kbits/s) comparada con el audio sin comprimir (1,41 Mbits/s; en el caso de audio muestreado a 44,1 KHz y resolución de 16 bits estéreo) es realmente buena. Tengamos en cuenta que la relación de compresión en este caso es de 11:1 (91%). Pero si prestamos atención realizando la escucha en un buen lugar y con un equipo de calidad media o alta, comienzan a manifestarse efectos o fenómenos que no son agradables, de hecho, si procedemos a elaborar una comparación entre el material sin comprimir y el compactado las diferencias resultarán apreciables. Por supuesto que para fines como escuchar música en el auto, en un walkman, en los "parlantitos" de la computadora, tener una melodía de fondo en el equipo de casa mientras cocinamos, hacer back up de charlas, conferencias, clases, etc. el formato es indiscutiblemente la mejor opción; mas no lo es, para grabaciones profesionales (como muchos piensan) o para disfrutar en un buen ambiente con un gran equipo, de una magnífica pieza sinfónica (lo de las sinfonías sólo es gusto personal).

Resulta curioso saber de personas que usan éste formato (MPEG-1 layer 3) en producciones discográficas de carácter comercial. Si bien es cierto, que para algunos efectos especiales (pájaros, trenes, ruido ambiente, etc.) la calidad resulta prácticamente indistinguible, me parece poco ético incorporar material de sonido que ha sido "cercenado" en su contenido esencial, solamente por no grabar o conseguir audio de calidad.

Insisto, la norma es realmente buena, pero, con las limitaciones que cada uno pueda tolerar. En lo personal, no uso el formato para oír o guardar música, aunque reconozco que es substancialmente mejor que un casete; y es un gran aliado cuando por correo electrónico tengo necesidad de enviar alguna "idea sonora". En los días que corren, son muy pocas las personas que "escuchan" música (me refiero a sentarse en un buen lugar con un equipo medianamente bueno y disfrutar del arte, como un buen melómano), generalmente se usa como "ruido de fondo agradable", por lo tanto se ha incrementado la utilización del MP3 notablemente (además de los costos insignificantes). Queda en cada uno verificar si el ahorro de dinero y bits nos beneficia, o entorpece nuestro sentido de la audición; y así poder decidir (tal vez sin ser tan extremista) si el MP3 es "magia o escoria".

Espero que el artículo haya servido para evacuar interrogantes en la materia. Si no es así, al menos deseo haber incrementado dudas para bien, o sea, que si algún lector se vio interesado por las cuestiones que traté superficialmente en esta nota y siente ansias por investigarlas con mayor profundidad, me siento halagado.

Para todos aquellos que quieran ahondar conocimientos expongo la bibliografía consultada.

Bibliografía:

- Overview of MPEG Audio: Current and Future Standards for Low Bit-Rate Audio Coding. Karlheinz Brandenburg; Marina Bosi
Audio Engineering Society 1997 Volume: 45 Issue: 1/2.
- Lossless Coding for Audio Discs. Peter Craven; Michael Gerzon
Audio Engineering Society 1996 Volume: 44 Issue: 9.
- Compatibility Matrixing of Multichannel Bit-Rate-Reduced Audio Signals

- Warner R. Th. Ten Kate. Audio Engineering Society 1996 Volume: 44 Issue: 12
- Audio Compression for Network Transmission (F). Martin Dietz; Harald Popp; Karlheinz Brandenburg; Robert Friedrich. Audio Engineering Society 1996 Volume: 44 Issue: 1/2.
 - A Bit Rate Scalable Perceptual Coder for MPEG-4 Audio. Bernhard Grill. Audio Engineering Society 1997 Preprint No: 4620 Convention No: 103.
 - Human Factors and the Acoustic Ecology: Considerations for Multimedia Audio Design. Bruce N. Walker; Gregory Kramer. Audio Engineering Society. 1996 Preprint No: 4343 Convention No: 101.
 - Audio Compression for Network Transmission. Martin Dietz; Harald Papp; Karlheinz Brandenburg; Robert Friedrich. Audio Engineering Society. 1995 Preprint No: 4129 Convention No: 99.
 - An Objective Method for Sound-quality Estimation of Compression Systems Michael Bank; Arie Taicher; Yakov Karabelnik. Audio Engineering Society. 1996 Preprint No: 4373 Convention No: 101.
 - Digital Audio Coding Based on the Look-Up Table. Do Hyung Kim; Yang Seock Seo. Audio Engineering Society 1995 Preprint No: 4091 Convention No: 99.
 - Combining MPEG, PCMCIA, ISDN, and GPS for Professional Audio Applications. Markus Erne. Audio Engineering Society. 1995 Preprint No: 4070 Convention No: 99.
 - ISO/IEC MPEG-2 Advanced Audio Coding. Marina Bosi; Karlheinz Brandenburg; Schuyler Quackenbush; Louis Fielder; Kenzo Akagiri; Hendrik Fuchs; Martin Dietz. Audio Engineering Society. 1996 Preprint No: 4382 Convention No: 101.
 - MPEG-2 NBC Audio—Stereo and Multichannel Coding Methods. James D. Johnston; Jürgen Herre; Mark Davis; Uwe Gbur. Audio Engineering Society. 1996 Preprint No: 4383 Convention No: 101.
 - Lossless Audio Signal Processing. James A. S. Angus Audio Engineering Society. 1997 Preprint No: 4632 Convention No: 103.
 - The Bewildering Wilderness. Dennis A. Bohn. Sound & Video Contractor. Volume 18 N° 11.
 - Elements of Acoustical Engineering. Harry Olson. Ed: D. Van Nostrand Company. Inc.
 - Music Speech Audio. William Strong; George R. Plitnik. Ed: Sound Print.
 - La Historia de un Pionero del Audio Digital. La Nación. Ricardo Sametband. 2001 Suplemento de informática del 5 de febrero.

Carlos Indio Gauvron
 (Docente F. I. U. B. A.)
 in_dio_ar@yahoo.com.ar