

Grabación digital: ¿DAT o MiniDisc?

Federico Miyara

El dilema

Existen en el mercado diversas tecnologías para el registro digital del sonido, y la mayoría de ellas proporciona, a primera vista, resultados aceptables. Sin embargo, existen diferencias importantes que pueden incidir en la calidad lograda, y por lo tanto es preciso conocer los alcances y las limitaciones de cada tecnología, para seleccionar la que mejor se adapta a nuestra ecuación costo-prestaciones.

Un poco de historia

La grabación digital del sonido se conoce desde hace varias décadas. De hecho, los primeros grabadores digitales se construyeron en 1967 utilizando la tecnología de cabezales rotativos de barrido helicoidal desarrollada para la grabación de video analógico. En 1982 aparecen en el mercado los primeros reproductores de discos compactos (CD) con tecnología óptica, cuyo prototipo había sido presentado en Europa en 1979. Esto marca un verdadero hito histórico: la posibilidad de acercar al consumo masivo el audio digital y sus virtudes (inalterabilidad, fidelidad, bajo ruido). Faltaba, sin embargo, un medio que permitiera no sólo reproducir sino también grabar digitalmente. Esta idea se lleva a la práctica en 1986, cuando se consigue adaptar y miniaturizar el cabezal rotativo, obteniéndose el formato R-DAT (rotary-head digital audio tape), más conocido como DAT. Pese a su portabilidad, la complejidad mecánica de los grabadores DAT (lo cual incide en su costo) y su acceso inherentemente secuencial han dificultado su adopción masiva, restringiéndose a un selecto conjunto de usuarios en el campo profesional o semiprofesional.

Este espacio pretendió ser cubierto por dos tecnologías surgidas contemporáneamente en 1991: el Cassette Digital Compact (DCC) y el MiniDisc (MD). El DCC utilizaba un cabezal fijo multipolar y grababa en cinta magnética. Nuevamente un formato de acceso secuencial, tuvo una vida efímera, pese a las expectativas de que llegaría a reemplazar al tradicional Cassette Compacto (CC) analógico, del cual era sucesor. El MD, en cambio, incorporaba por primera vez un diminuto disco regrabable con tecnología óptica en buena parte heredada del CD. Esta tecnología, además de permitir el acceso aleatorio, permitía opciones avanzadas como la posibilidad de registrar y editar información administrativa (por ejemplo el título de las pistas), una fragmentación similar a la de los discos de computadora (útil para optimizar el espacio liberado al borrar pistas) y una gran robustez frente a vibraciones o movimientos bruscos merced a la utilización de una memoria buffer. Todo esto se combinaba con una mayor simplicidad mecánica, lo cual permitía un costo muy competitivo (si bien nunca comparable al del CC).

Es interesante el hecho de que ambas tecnologías requerían el uso de técnicas de compresión de datos, ya que debían almacenar un tiempo equivalente al de un CD (74 minutos) en un espacio mucho menor. Para el DCC se desarrolló la compresión PASC (Precision-adaptive Sub-band Coding), similar al MPEG-1 nivel 1 (precursor del esquema de compresión MPEG-1 nivel 3, hoy popularizado como MP3), que lograba una

compresión 4:1, en tanto el MD adoptó la compresión ATRAC (Adaptive Transform Acoustic Coding), con una reducción de datos de 5:1. Ambas compresiones se basan en el fenómeno psicoacústico del enmascaramiento.

Últimamente se han incorporado al mercado grabadores digitales portátiles en disco rígido que equiparan al DAT en lo que respecta a características, pero requieren la descarga a una computadora (lo cual se logra a través de un puerto USB) ya que su capacidad de almacenamiento es limitada (por ejemplo, 5 Gb).

Rudimentos de la grabación digital

Cualquiera sea la tecnología utilizada para la grabación digital, la señal de audio debe someterse primero a varios procesos. En primer lugar, debe ser muestreada, es decir que deben tomarse muestras de la señal analógica a intervalos regulares, a razón de f_s muestras por segundo, donde f_s es la tasa (o frecuencia) de muestreo (figura 1).

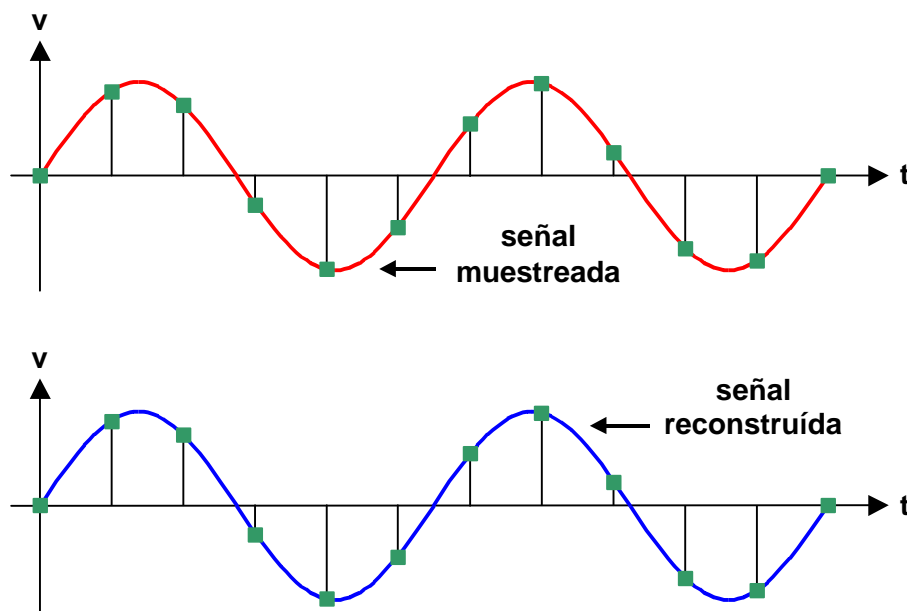


Figura 1. Proceso de muestreo y reconstrucción

La cantidad de muestras por segundo debe cumplir con una condición llamada Teorema del Muestreo, que establece que para que la señal muestreada pueda luego recuperarse sin distorsiones la tasa de muestreo debe ser por lo menos el doble de la máxima frecuencia contenida en la señal, independientemente de que dicha frecuencia sea útil o sólo ruido. Si no se cumple esta condición, se produce un tipo de distorsión conocido como “aliasing” (figura 2).

Para evitar esta situación, todos los sistemas de muestreo aplican, antes del muestreo, un filtro, conocido como filtro antialias, que limita las frecuencias superiores a lo estrictamente necesario. Para el caso del audio de alta fidelidad, esto corresponde a un máximo de 20 kHz. La tasa de muestreo debe ser entonces mayor de 40 kHz, y por razones prácticas, se ha adoptado una tasa de muestreo normalizada de 44,1 kHz.

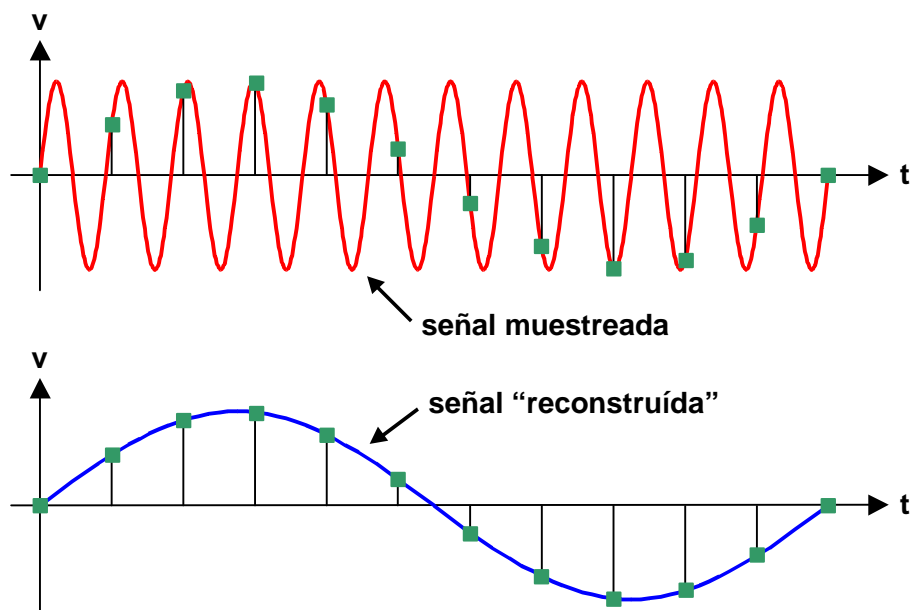


Figura 2. Fenómeno de aliasing con reconstrucción distorsionada

Una vez realizado el muestreo, se debe proceder a la digitalización, es decir, la conversión de la señal analógica en números, expresados con una cantidad de dígitos binarios o bits denominada resolución (figura 3).

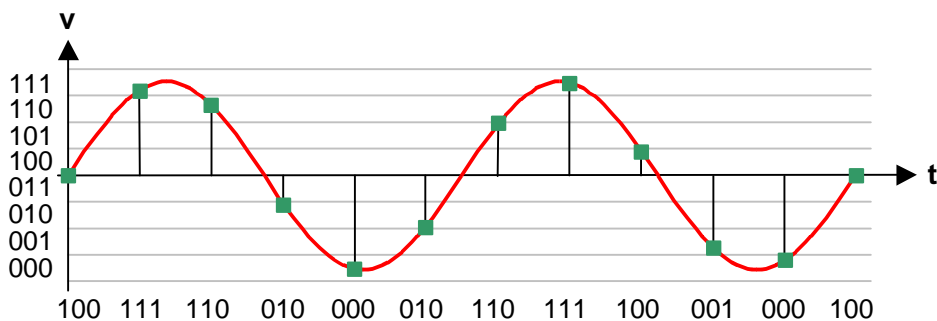


Figura 3. Digitalización de la onda muestreada

La digitalización es, de hecho, una aproximación, ya que se reemplazan valores que varían en forma continua por valores discretos. Así, todos los posibles valores entre el 101 (es decir 5 en notación binaria) y el 110 (6 en notación binaria) se reemplazan por 101. Evidentemente, cuanto mayor sea la cantidad de bits utilizada, mayor será la exactitud de la representación.

Al ser la señal digitalizada diferente de la real, podemos interpretar que se ha “contaminado” con ruido; un tipo de ruido denominado ruido de digitalización. Este ruido es tanto menor cuanto mayor sea la resolución en bits. El parámetro que mide la incidencia del ruido en un sistema es la *relación señal/ruido*, es decir la diferencia en decibelios entre la máxima señal y el ruido. En el caso del ruido de digitalización, la re-

lación señal/ruido máxima alcanzable con n bits (suponiendo que el resto del sistema sea perfecto, es decir que no agregue más ruido) puede aproximarse mediante la fórmula

$$S/R_{\text{máx}} \cong 6 \cdot n \quad [\text{dB}].$$

Así, con 16 bits (la resolución estándar del CD, el DAT y el MD) es posible en teoría lograr una relación señal/ruido de alrededor de 96 dB. En la práctica esta marca suele *no* alcanzarse debido a imperfecciones de las diversas etapas análogas (preamplificadores, conversores, filtros antialias, etc.).

La grabación digital consiste simplemente en almacenar los valores de las sucesivas muestras digitalizadas en alguna forma de memoria masiva (ya que es una gran cantidad de muestras).

Tanto en las cintas como en los discos, dado que la densidad de información es muy alta, existe cierta probabilidad de errores derivados de imperfecciones o deterioros en la superficie. Para prevenir esto, se introduce un código que permite la detección y corrección de errores (código de Reed-Solomon), lo cual implica incrementar la cantidad de datos grabados. Los códigos correctores de errores se basan en la redundancia.

La información perdida

Todo sistema de muestreo y conversión analógico-digital implica la pérdida de alguna información originalmente existente en la señal. Así, el muestreo significa que de la variación continua de la señal se eligen valores o muestras a intervalos regulares, ignorándose lo que sucede entre dos muestras sucesivas. Cuando esto se combina con el filtro antialias que precede al proceso de muestreo, el resultado es una pérdida o recorte de información de alta frecuencia.

Análogamente, la digitalización equivale a renunciar al valor exacto de la muestra y a reemplazarlo por un valor apenas aproximado, cuya única virtud reside en que se puede representar con los n bits elegidos (por ejemplo, $n = 16$). En este caso, se están perdiendo las variaciones más pequeñas de la señal, es decir, todas las que quedan comprendidas entre dos escalones o cifras consecutivas. Esto implica una pérdida de rango dinámico.

Para que el sistema de muestreo y digitalización sea satisfactorio, es preciso que la información perdida sea irrelevante, es decir, inaudible. Ello se logra eligiendo la frecuencia de muestreo de modo tal que el ancho de banda resultante incluya todas las señales audibles, y adoptando un número de bits que asegure un rango dinámico igual o superior al del oído.

Cinta de audio digital (DAT)

Este formato consiste en un cassette de cinta magnética de 3,81 mm de ancho en el cual se graban delgadas pistas paralelas de 13,6 μm inclinadas unos 6° respecto al borde de la cinta que contienen la señal digitalizada y codificada como PCM (pulse code modulation). La codificación PCM consiste básicamente en una sucesión en el tiempo de valores bajos y altos de una señal eléctrica (u óptica, o magnética) representando los sucesivos 0's y 1's de un número binario (en este caso proveniente de la conversión analógico-digital de una muestra analógica del sonido). Esta codificación no introduce pérdida de información ninguna respecto a la señal ya muestreada y digitalizada.

MiniDisc (MD)

El MiniDisc es un pequeño disco de 64 mm de diámetro dentro de una cubierta protectora. La grabación se efectúa mediante un haz láser que calienta a 180 °C una pequeña área del disco, mientras una bobina la magnetiza con una u otra polaridad (lo cual sólo es posible a esa temperatura, denominada temperatura de Curie). La lectura se logra mediante un efecto magneto-óptico denominado efecto Kerr, por el cual la presencia de un campo magnético produce un cambio de la polarización del haz láser.

Debido a que la densidad de información (es decir la cantidad de datos por unidad de superficie —que está determinada principalmente por la longitud de onda del láser utilizado, 780 nm—) es igual a la del CD, pero su superficie bastante menor, el MD necesita incorporar técnicas de compresión (reducción) de datos. Para comprender las posibles limitaciones de este sistema, conviene inspeccionar la naturaleza del proceso de compresión, llamado genéricamente ATRAC (Adaptive Transform Acoustic Coding).

La compresión ATRAC

El sistema se basa en el fenómeno psicoacústico de enmascaramiento, consistente en que un sonido puede volverse inaudible en presencia de otro. En ausencia de otros sonidos, cada frecuencia f tiene asociado un umbral de audición, es decir, un nivel de presión sonora por debajo del cual el sonido no se percibe (curva roja en la figura 4). Si ahora se hace sonar otro sonido simultáneamente con f (por ejemplo una banda angosta de ruido centrada en 1 kHz), se observa que el umbral de audición correspondiente a la frecuencia f aumenta (curva azul en la figura 4), es decir que se requiere mayor nivel para que f se alcance a escuchar. La presencia del otro sonido ha *enmascarado* a f . En el ejemplo, el umbral para la frecuencia f ha aumentado de 4 dB a 28 dB, es decir 24 dB (el aumento se indica con la línea verde en la figura 4). Esto significa, en particular, que puede tolerarse un ruido 24 dB más intenso, ya que el mismo resultará inaudible.

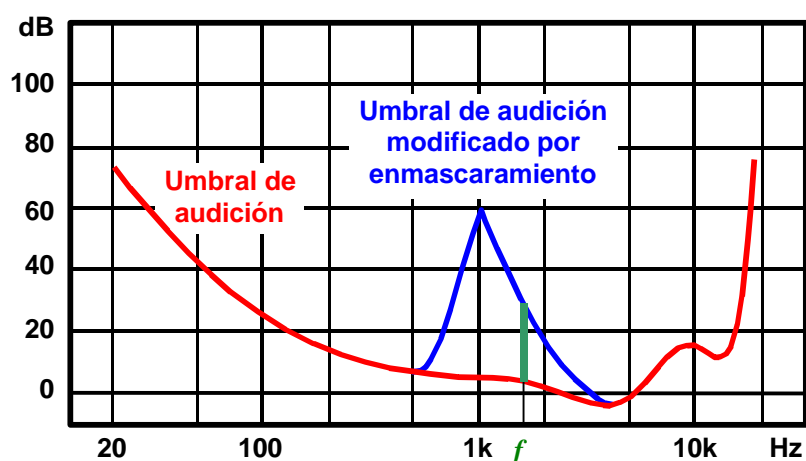


Figura 4. Umbral de audición y su modificación por el enmascaramiento producido por una banda de ruido centrada en 1 kHz.

El codificador ATRAC aprovecha este fenómeno del siguiente modo. Dado que el ruido admisible aumentó 24 dB, la relación señal/ruido necesaria disminuye 24 dB, por lo cual se requieren $24/6 = 4$ bits menos. En lugar de 16 bits se requerirán, por lo tanto, 12 bits.

El ejemplo anterior era un caso simple, en el cual el sonido enmascarante era una banda de ruido muy angosta. En un caso real (por ejemplo un fragmento musical), tendremos generalmente un espectro complejo y variable en el tiempo. Un espectro complejo significa muchos sonidos simultáneos enmascarándose mutuamente. El codificador actúa descomponiendo primero la señal en cierto número de bandas por medio de un banco de filtros, y luego determinando el efecto enmascarante sobre cada banda de las otras bandas. Este efecto enmascarante se manifiesta en general como un importante aumento del umbral, lo cual significa la posibilidad de asignar una cantidad de bits mucho menor que 16 a cada una de las bandas.

Dado que las características espectrales de la señal son variables en el tiempo, la cantidad de bits asignada se actualiza cada 11,6 milisegundos, excepto en los sonidos de variación muy rápida (ataques), en que la actualización puede realizarse cada 2,9 milisegundos para las frecuencias bajas y medias, y cada 1,45 milisegundos para las altas frecuencias. El codificador guarda en el propio disco dos parámetros: la longitud de la palabra (es decir la cantidad de bits asignados a cada muestra¹) y el *factor de escala*, que indica el tamaño que tiene cada bit.

Una característica interesante del ATRAC es que el decodificador es independiente del codificador. Esto es, al guardar información suficiente para reconstruir la señal, el decodificador es muy simple y completamente estándar. No sucede lo mismo con el codificador, cuyo algoritmo puede ir evolucionando con el tiempo, a medida que se van perfeccionando las técnicas de compresión, o que mejora la precisión aritmética de los procesadores. Ello ha dado origen a sucesivas versiones del codificador ATRAC, que dan resultados cada vez mejores, pero que, sin embargo, pueden ser decodificadas por un mismo decodificador. La versión actual es la 4.5.

Efectos residuales de la compresión

Al considerar la posibilidad de utilizar el MiniDisc para registros digitales de sonido, es muy importante tener en cuenta el efecto residual de la compresión. Debido a que el enmascaramiento es un fenómeno espectralmente local, es decir que un sonido enmascara primordialmente a las frecuencias vecinas (figura 4), cabe esperar que el peor efecto se produzca en los casos de gran cantidad de frecuencias muy cercanas.

Para evaluar estos efectos, se ensayó un grabador MiniDisc portátil (el MZ-R70 de Sony). Los ensayos se hicieron utilizando tres tipos de señales generadas por computadora, enviadas desde la salida de línea de la placa de sonido, grabadas en el grabador de MiniDisc, reproducidas por éste y reingresadas a la computadora a través de la entrada de línea de la placa. La primera señal estuvo constituida por un barrido de tonos puros correspondientes a las frecuencias centrales de las bandas de tercio de octava desde 20 Hz hasta 20 kHz. En este caso los tonos se presentan en forma aislada, por lo cual no existen efectos de enmascaramiento. El sistema exhibe un énfasis de hasta 5 dB en baja frecuencia y una atenuación de casi 7 dB en 20 kHz (figura 5). Por consiguiente, es necesaria una ecualización, por ejemplo a través de un filtro FFT, si se desea una respuesta verdaderamente plana. Al parecer, este énfasis en los graves es para compensar

¹ En realidad no se guardan las muestras sino los coeficientes espectrales.

la pobre respuesta en baja frecuencia de la mayoría de las vías acústicas de mediana calidad.

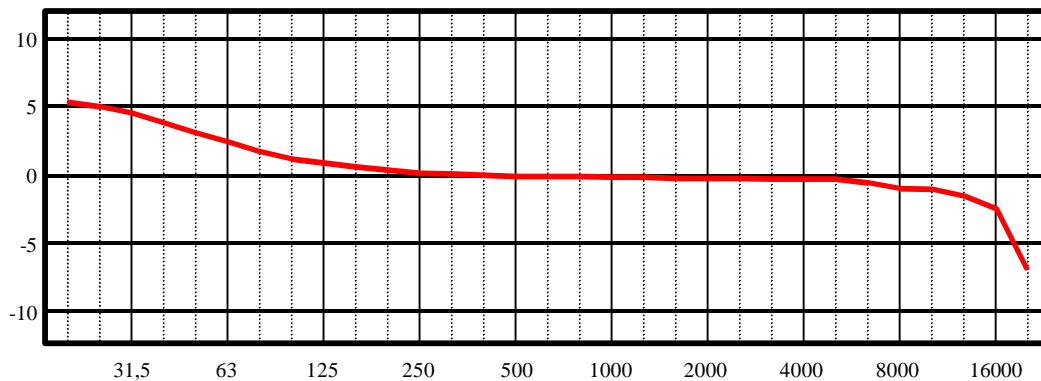


Figura 5. Respuesta en frecuencia del MiniDisc ante tonos puros de frecuencias iguales a los centros de las bandas de tercio de octava

La segunda señal utilizada fue un tono de 100 Hz y 200 armónicos de igual amplitud, todos en fase. En este caso se pone a prueba el algoritmo de adaptación de la precisión según el enmascaramiento, ya que los tonos utilizados tienen una tendencia a enmascarar a los vecinos, reduciendo el rango dinámico disponible. Se observa en este caso una reducción brusca de unos 4 dB en las proximidades de 15 kHz. Hasta donde pudo observarse, los armónicos mantienen en general la fase de manera que la forma de onda rescatada es similar a la original.

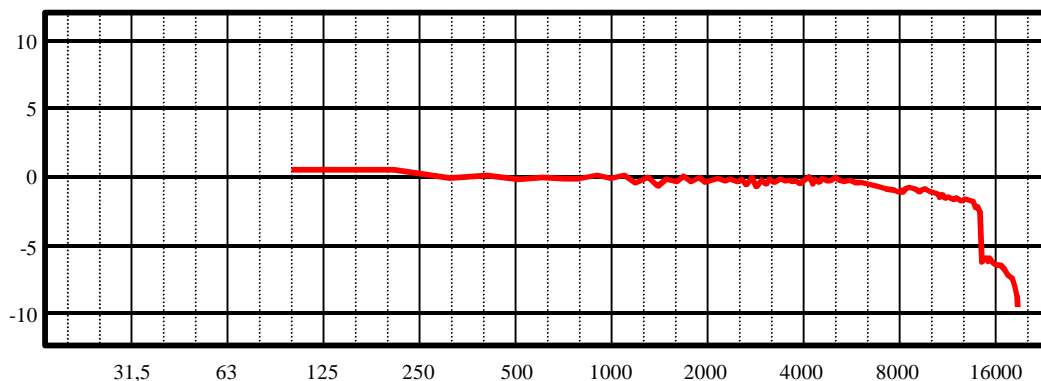


Figura 6. Respuesta en frecuencia del MiniDisc ante un sonido formado por 200 armónicos de 100 Hz en fase

La tercera señal correspondió un fondo de ruido blanco al cual se le superpone un tono de 1 kHz que va adoptando niveles decrecientes de 5 en 5 dB, comenzando por un nivel igual al del ruido blanco, hasta quedar sumergido en el ruido. En este ensayo se pretendió verificar hasta cuándo era todavía posible rescatar el tono, teniendo en cuenta el enmascaramiento causado por el ruido espectralmente vecino. El ensayo se realizó de dos formas diferentes. En la primera, el ruido era perfectamente blanco, y en la segunda

se le eliminó, por filtrado FFT, una pequeña banda alrededor de 1 kHz. En ambos casos se agrega posteriormente el tono de 1 kHz de nivel decreciente. En el primer caso, el tono puede ser rescatado por medio de un filtro pasabanda angosto alrededor de 1 kHz hasta que se vuelve unos 25 dB inferior al nivel del ruido blanco. Para valores más bajos, no es posible separar el ruido filtrado del tono. En el segundo caso, la reducción llega fácilmente a 43 dB por debajo del nivel del ruido blanco.

El mismo test, aplicado a la señal digital originalmente generada por la computadora, arroja valores de 30 dB y 59 dB, por debajo del nivel del ruido blanco, o blanco con 1 kHz suprimido, respectivamente. Resultados similares pueden apreciarse para el DAT, habiéndose utilizado un equipo portátil PCM-M1, también de Sony.

Conclusión

Los ensayos revelan algunas diferencias objetivas entre las respuestas del DAT y del MiniDisc. La respuesta del MiniDisc a señales simples (tonos puros) es casi indistinguible de la de un DAT. Para señales espectralmente complejas (gran cantidad de armónicos o ruidos de bandas), se produce una alteración equivalente a un aumento del ruido de fondo que resulta prácticamente inaudible (por la propia naturaleza psicoacústica de la compresión), pero que está presente y podría manifestarse en situaciones en las que la señal grabada deba ser sometida a procesamientos que involucren ecualizaciones complejas o filtrados importantes.

En aquellos casos en que se esté grabando una señal que no experimentará modificaciones ulteriores muy grandes, la grabación en MiniDisc será perfectamente aceptable.

Finalmente, cabe acotar que las comparaciones fueron efectuadas sobre equipos portátiles. En los equipos de mesa, los conversores suelen ser de mejor calidad, y debido al proceso de compresión, en algunos casos se anuncian relaciones señal/ruido mayores de 96 dB (por ejemplo 100 dB), lo cual en el DAT no es posible por ser un formato perfectamente lineal (y como tal debe atenerse a los 96 dB teóricos).

LECTURAS COMPLEMENTARIAS

1. Federico Miyara: “¿Ruido o señal? La otra información. En defensa del registro digital del ruido urbano”. 4ta Jornada Regional sobre Ruido Urbano. Montevideo, 14/7/01
2. Watkinson, John: “Audio Digital”. Editorial Paraninfo. Madrid, 1996.
3. Kyoya Tsutsui, Hiroshi Suzuki, Osamu Shimoyoshi, Mito Sonohara, Kenzo Akagiri, Robert M. Heddle: “ATRAC: Adaptive Transform Acoustic Coding for MiniDisc”. Sony Corporate Research Laboratories, 6-7-35 Kitashinagawa, Shinagawa-ku, Tokyo 141 Japan. Reprinted from the 93rd Audio Engineering Society Convention in San Francisco, 1992 October 1-4. Internet: http://www.minidisc.org/aes_atrac.html
4. Tadao Yoshida: “The Rewritable MiniDisc System”. Advanced Development Laboratories. Sony Corporation, Tokyo, Japan. Reprinted from the Proceedings of the IEEE, USA vol. 82 no. 10 pp. 1492-1500, OCT 1994. Internet: http://www.minidisc.org/ieee_paper.html