

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 385 986**

51 Int. Cl.:
G11B 20/00 (2006.01)
G10H 1/36 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **05009434 .1**
96 Fecha de presentación: **29.04.2005**
97 Número de publicación de la solicitud: **1592008**
97 Fecha de publicación de la solicitud: **02.11.2005**

54 Título: **Grabación en estéreo compatible con multicanal**

30 Prioridad:
30.04.2004 US 567041 P

45 Fecha de publicación de la mención BOPI:
06.08.2012

45 Fecha de la publicación del folleto de la patente:
06.08.2012

73 Titular/es:
AURO TECHNOLOGIES NV
Kievitstraat 42
2400 Mol , BE

72 Inventor/es:
Van den Berghe, Guido

74 Agente/Representante:
de Elzaburu Márquez, Alberto

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

ES 2 385 986 T3

DESCRIPCIÓN

Grabación en estéreo compatible con multicanal

Campo técnico de la invención

5 La presente invención se refiere a las técnicas de formateo para el almacenamiento de datos, por ejemplo para dispositivos de estado sólido o disco óptico, tal como DVD (Disco Versátil Digital). Más particularmente, la presente invención se describirá con referencia al formato de CD de audio, un formato de electrónica de consumo para música pregrabada en disco compacto, de CD+G(w), pero no está limitada a esta forma de grabación. Este formato de CD de audio puede ser, por ejemplo, CD-DA (Compact Disc Digital Audio, el formato de CD de música original, que almacena la información de audio como datos digitales PCM modulados por codificación de impulsos) compatibles con Redbook, y
10 puede usar el modo CD+G usuario para almacenar datos adicionales. CD+G(w) es una técnica de codificación y decodificación útil para, pero no que se limita a, funciones de Cantar y Tocar en grupo (Sing & Play-Along), tales como karaoke y la reproducción automatizada de instrumentos, usando grabación en estéreo compatible con multicanal.

Antecedentes de la invención

15 Generalmente, el CD+G existente es un formato de CD de audio, que tiene capacidades limitadas para almacenar datos gráficos adicionales y se usa, principalmente, en aplicaciones de karaoke. La parte de audio de CD+G puede ser reproducida en cualquier CD de audio estándar. Para la visualización de gráficos, se requiere un reproductor de CDG, un reproductor de CD especial con salida de TV, por ejemplo, un reproductor de CD de karaoke. En el formato CD+G, los datos de audio son grabados en un área de programa, mientras que los datos adicionales pueden ser almacenados en una zona de gestión de programas, la cual está dividida en ocho canales sub-código, los canales P a W. El canal P indica
20 el inicio y el final de cada pista, alternando entre 0 y 1. El canal Q contiene los códigos de tiempo y una tabla de contenidos. Más datos adicionales contenidos en los canales de sub-código R a W, pueden permitir que se visualicen gráficos y texto mientras se está reproduciendo música, tal como se usa, por ejemplo, para representar imágenes fijas o representar letras, por ejemplo, para un karaoke.

El estándar CD+G define diferentes modos de CD+G, que son los modos ZERO, GRAPHICS (karaoke), MIDI y USER:

- 25 • El modo ZERO es un modo en el que no se realiza ninguna operación en una pantalla de visualización.
- El modo GRAPHICS permite representar imágenes o símbolos, por ejemplo, para un karaoke.
- El modo MIDI proporciona un canal de datos de 3.125 kb/s máximo para los datos de Musical Instrument Digital Interface (MIDI), tal como especifica la Asociación Internacional MIDI.
- 30 • El modo USER está destinado para aplicaciones profesionales. El significado de los datos es específico de la aplicación.

El documento US-5.852.800 describe un procedimiento y un aparato para la modulación y el mezclado, controlados por usuario, de datos comprimidos, almacenados digitalmente. En este documento, la pluralidad de canales son almacenados por separado en un CD ROM. De esta manera, la pista de audio en el CD ROM corresponde a la pluralidad de canales de audio, mono, independientes. Esta pluralidad de canales son recuperados desde el CD ROM, son
35 modulados y mezclados según las direcciones y las órdenes recibidas del usuario, y se crea una señal de suma para su presentación a una audiencia. La señal de suma se obtiene descomprimiendo los canales de audio por separado, multiplicando cada canal de audio por los valores de intensidad correctos y, a continuación, sumando entre sí todos los canales seleccionados. Cada canal puede ser enmascarado o silenciado por completo.

40 En el documento EP-0757506, los canales de audio separados son codificados como paquetes de datos específicos en un flujo de bits multicanal y, como tales, son accesibles por el decodificador. A nivel del decodificador, el usuario tiene control sobre el modo downmix (conversión de multicanal a estéreo) para activar o desactivar ciertos canales de audio en la conversión de multicanal a estéreo.

El documento US-6405163 describe el uso de características de frecuencia presentes igualmente en ambos canales izquierdo y derecho (desde el punto de vista de las características de frecuencia) para "reducir" o "amplificar" la información de audio que está centrada. Esta solución permite la "atenuación" de la información de audio central.

Resumen de la invención

Un objeto de la presente invención es proporcionar un sistema de audio mejorado.

El objetivo anterior se consigue mediante un codificador según la reivindicación 1 y decodificador según la reivindicación 12.

Los aspectos particulares y preferentes de la invención se exponen en las reivindicaciones independientes y dependientes adjuntas. Las características de las reivindicaciones dependientes pueden ser combinadas con las características de las reivindicaciones independientes y con características de otras reivindicaciones dependientes, según sea apropiado, y no sólo como se expone explícitamente en las reivindicaciones.

- 5 El codificador de la presente invención puede ser usado para mezclar una pluralidad de canales independientes de audio mono en una grabación en estéreo y generar un conjunto restringido de parámetros adicionales usados para la masterización de una pista de audio de un dispositivo de almacenamiento, estando construidos la pluralidad de canales independientes de audio mono de manera que el dispositivo de almacenamiento puede ser reproducido usando un reproductor de disco óptico de manera que, en un primer modo, la totalidad de la pluralidad de canales independientes de audio mono son reproducidos como la grabación en estéreo y, en un segundo modo, al menos uno de entre la pluralidad de canales independientes de audio mono puede ser des-mezclado y la grabación en estéreo puede ser reproducida con al menos un canal de audio mono eliminado.

El dispositivo de almacenamiento puede ser un disco óptico, tal como, por ejemplo, un disco de CD de audio digital.

- 15 En un codificador según una realización de la presente invención, preferentemente, el conjunto restringido de parámetros adicionales no exceden el tamaño de datos disponible como datos de sub-código dentro del modo CD+G USER.

- 20 El decodificador de la presente invención puede ser usado para reconstruir, en un primer modo, una pluralidad de canales independientes de audio mono en una grabación en estéreo usando un conjunto restringido de parámetros adicionales usados para la masterización de una pista de audio de un dispositivo de almacenamiento, cuando se lee únicamente la grabación de audio en estéreo y los parámetros adicionales desde el dispositivo de almacenamiento y, en un segundo modo, al menos uno de entre la pluralidad de canales independientes de audio mono puede ser des-mezclado y la grabación en estéreo es reproducida con el al menos un canal de audio eliminado y no reproducido.

Los canales reconstruidos pueden estar no correlacionados, y pueden ser sustancialmente idénticos, perceptualmente, a los canales mono originales.

- 25 En un decodificador según una realización de la presente invención, el conjunto restringido de parámetros adicionales y los datos de mezclado, de desplazamiento lateral y/o atenuación asociados pueden proporcionar información suficiente al decodificador para aislar, eliminar o atenuar un único canal o más canales durante la reproducción en tiempo real.

En un tercer aspecto, la presente invención proporciona un sistema de audio según la reivindicación 21.

- 30 En un sistema de audio según una realización de la presente invención, el sistema soporta aplicaciones o configuraciones de audio diferentes, incluyendo filtrado de solista (estéreo), filtrado de voz solista (mono) e instrumento solista (mono), con reproducción de audio cuadrafónico verdadero.

- 35 Además, cuando se usa en una aplicación de karaoke, la información de las letras puede estar incluida en el dispositivo de almacenamiento, en un formato tal que puede ser visualizada en una pantalla de visualización. Cuando se usa con una aplicación de reproductor de instrumento, el instrumento principal puede ser filtrado del audio y puede ser remplazado por un reproductor de instrumento controlado, donde los datos de control son una parte de los datos sub-código en el dispositivo de almacenamiento. Puede proporcionarse una combinación de la aplicación de Karaoke y la aplicación de reproductor de instrumento.

También, pueden incluirse, en el dispositivo de almacenamiento, datos adicionales del filtro de audio para mejorar la expansión temporal y/o el cambio de tono.

- 40 Cuando se usa en aplicaciones de audio de sonido envolvente, el decodificador puede ser capaz de reconstruir los canales de audio estéreo traseros totalmente aislados de los canales estéreo frontales, a partir de un dispositivo de almacenamiento con codificación cuadrafónica verdadera. Cuando se usa en aplicaciones de sonido envolvente, el decodificador puede ser capaz de panoramizar dinámicamente la voz principal (mono) y/o el instrumento principal (mono) a partir de un disco óptico de un dispositivo de almacenamiento codificado para Karaoke, en el espacio auditivo 3D, para reproducir el efecto de una representación en vivo, cuando el artista se está moviendo en el escenario.

- 45 Los datos MIDI pueden estar integrados en el dispositivo de almacenamiento. El sistema de audio puede tener una precisión de temporización 272 μ s por evento MIDI, así como una resolución de velocidad MIDI de 256 niveles diferentes.

- 50 En un sistema de audio según las realizaciones de la presente invención, el decodificador puede tener medios para ejecutar la operación inversa del codificador y tanto el decodificador como el codificador pueden ser implementados en aritmética entera de 32 bits, siendo independientes de la implementación en coma flotante. En primer lugar, el codificador puede generar canales de audio aproximados antes de mezclar estos canales, usando parámetros del filtro de audio generados como valores en coma flotante primero, que son el resultado de una optimización con objetivos configurables

por el usuario, que pueden incluir características perceptuales de audio. El sistema de audio puede comprender además la conversión a aritmética entera, teniendo el codificador y/o el decodificador medios para compensar los errores de redondeo y de conversión para proporcionar un procedimiento de decodificación que es la operación inversa de la codificación.

5 En un sistema de audio según las realizaciones de la presente invención, los parámetros del filtro de audio requeridos por el decodificador y que son una parte del conjunto de parámetros de datos, pueden ser regenerados para cada período de tiempo, que es corto en comparación con la resolución de audio del oído humano. El período de tiempo puede ser de 40 milisegundos de audio.

10 En un sistema de audio según las realizaciones de la presente invención, el codificador puede incorporar un compresor dinámico que previene rebasamientos durante el mezclado y/o la codificación de la pluralidad de canales mono en un canal estéreo convertido desde multicanal.

El formato CD+G(w), que puede ser usado con la presente invención, usa el modo USER del CD+G estándar para almacenar datos adicionales. Una parte de estos datos, la sección de datos de filtro, define diversos parámetros del filtro de audio. Otra parte, la sección de datos MIDI, define los datos usados para los dispositivos compatibles con MIDI. Las
15 letras de canciones para cantar en grupo pueden estar integradas dentro de la sección MIDI de los datos.

En un aspecto de la presente invención, la sección de datos del filtro de audio contiene los parámetros del filtro de audio que permiten a un decodificador CD+G(w) filtrar pistas individuales, tales como la voz principal y/o un instrumento solista de la grabación de audio en estéreo. Conforme la grabación combinada de audio en estéreo es filtrada, la cantidad de
20 datos adicionales necesarios para almacenar la grabación en estéreo no es significativamente mayor que los usados para almacenar la grabación en estéreo usando los parámetros estándar. La sección de datos MIDI contiene datos para controlar los dispositivos controlados digitalmente (MIDI), tales como pianos digitales o pianolas u otros instrumentos musicales controlados MIDI, equipo ligero controlador de MIDI o virtualmente cualquier equipo compatible con MIDI.

En otro aspecto de la presente invención, la sección de datos del filtro de audio es el resultado de un complejo esquema de codificación, así como una remasterización de las grabaciones de las pistas máster originales a partir de las grabaciones de audio máster. El decodificador CD+G(w) es capaz de aislar o eliminar 1 ó 2 o más "pistas principales" que
25 están mezcladas en el estéreo convertido desde multicanal, remasterizado, sin la necesidad de almacenar estas "pistas principales" separadas en el CD. En combinación con la salida en TV de las letras y dispositivos con capacidad MIDI, tales como, por ejemplo, instrumentos musicales automatizados, este formato CD+G(w) es un formato de CD único y proporciona solución a diversas características diferentes: tales como el karaoke o canto en grupo cuando la pista vocal principal es filtrada/eliminada de la mezcla en estéreo, o reproducción en grupo cuando una pista del instrumento principal es filtrada/eliminada, o incluso reproducción en grupo y canto en grupo, cuando las partes vocal principal e instrumento principal son filtradas/eliminadas y los datos MIDI son usados para mostrar letras, así como para controlar un reproductor de instrumento automatizado. Cuando se eliminan la voz principal y el instrumento principal, las voces/los instrumentos de
30 respaldo estéreo son todavía audibles, sin ningún artefacto. Junto con estas características, una grabación de audio compatible con CD-DA de audio cuadrafónico verdadero se convierte en una realidad, ya que la naturaleza básica de este formato y su grabación y esquema de masterización y codificación permiten grabar y convertir cuatro pistas independientes de audio mono en 1 grabación en estéreo, permitiendo, asimismo, una operación de des-mezclado.

En otro aspecto de la presente invención, se proporciona una técnica de grabación multicanal, preferentemente, una técnica de grabación de CD de audio multicanal, para grabar en un canal mezclado una pluralidad de canales mono, en el
40 que cada uno tiene una frecuencia de muestreo original, en el que, antes de la mezcla, cada canal mono es redefinido usando al menos la mitad de su frecuencia de muestreo original y un conjunto limitado de parámetros adicionales. La técnica de grabación según este aspecto de la presente invención permite, durante la reproducción de las pistas grabadas, obtener un 100% de filtrado de canal, es decir, en un primer modo, toda la pluralidad de canales independientes mono (audio) pueden ser reproducidos como una grabación estéreo y, en un segundo modo, al menos uno de entre la pluralidad de canales independientes mono (audio) canales puede ser des-mezclado y la grabación estéreo puede ser reproducida con al menos un canal mono (audio) eliminado. La frecuencia de muestreo de los canales mono redefinidos depende del número de canales a mezclar, es decir, es la mitad de la frecuencia de muestreo original de los canales mono si se van a
45 mezclar dos canales mono, o es una cuarta parte de la frecuencia de muestreo de canal mono original si se van a mezclar cuatro canales, etc. Solo como un ejemplo, si se van a mezclar dos canales mono de 16-bit con una frecuencia de muestreo de 44 kHz, según el presente aspecto de la invención, estos canales mono son redefinidos como dos canales mono de 15-bits con una frecuencia de muestreo de 22 kHz, y estos canales mono redefinidos son mezclados. La redefinición de los canales mono de 16-bits puede ser realizada tomando, desde el primer canal, sólo las muestras pares, y tomando, desde el segundo canal, sólo las muestras impares. Como otro ejemplo, si se van a mezclar cuatro canales mono de 16-bits, cada uno con una frecuencia de muestreo de 44 kHz, los canales mono pueden ser redefinidos como
50 cuatro canales de 14-bits, cada uno con una frecuencia de muestreo de 11 kHz, y estos cuatro canales mono redefinidos pueden ser mezclados. La redefinición de los canales mono de 16 bits puede realizarse tomando, del primer canal, la 4-

ésima muestra, desde el segundo canal, la $(4k + 1)$ -ésima muestra, desde el tercer canal, la $(4k + 2)$ -ésima muestra, y desde el cuarto canal, la $(4k + 3)$ -ésima muestra.

En un aspecto adicional de la presente invención, un canal mono casi equivalente puede ser obtenido a partir del canal mono definido y el conjunto limitado de parámetros adicionales, por ejemplo, interpolando las muestras que han sido descartadas durante el procedimiento de redefinición. Esto significa que, a partir de los canales mono redefinidos grabados, por ejemplo dos canales mono 15-bits con una frecuencia de muestreo de 22 kHz o cuatro canales mono de 14-bit con una frecuencia de muestreo de 11 kHz, junto con el conjunto de parámetros complementarios, pueden obtenerse canales mono de 16 bits, aproximados a los originales, que tienen una frecuencia de muestreo de 44 kHz.

Las características anteriores y otras características y funciones y ventajas de la presente invención se harán evidentes a partir de la descripción detallada siguiente, tomada en conjunción con los dibujos adjuntos, que ilustran, a modo de ejemplo, los principios de la invención. Esta descripción se proporciona solamente como ejemplo, sin limitar el alcance de la invención. Los números de referencia indicados más adelante se refieren a los dibujos adjuntos.

Breve descripción de los dibujos

La Fig. 1 ilustra un flujo de procesamiento para la codificación de audio para aplicaciones según una primera realización de la presente invención.

La Fig. 2 ilustra un flujo de procesamiento para la grabación de audio cuadrafónico verdadero según una segunda realización de la presente invención.

La Fig. 3 ilustra la creación de $CD+G(w)$.

La Fig. 4 ilustra un primer flujo A de audio (serie 1) y un flujo A' de audio aproximado (serie 2).

La Fig. 5 ilustra un segundo flujo B de audio (serie 1) y un flujo B' de audio aproximado (serie 2).

La Fig. 6 ilustra un flujo A+B de audio mezclado (serie 1) y un flujo A'+B' de audio mezclado aproximado (serie 2).

La Fig. 7 ilustra la aproximación de muestra de la muestra A_5 basada en un filtro FIR.

La Fig. 8 ilustra una compresión de audio debida a una mezcla o una interpolación.

La Fig. 9 es una tabla que ilustra una comparación de las propiedades de datos MIDI para los datos MIDI estándar y para los datos MIDI incluidos en el formato $CD+G(w)$ tal como se usa según las realizaciones de la presente invención.

Descripción de realizaciones ilustrativas

La presente invención se describirá con respecto a las realizaciones particulares y con referencia a ciertos dibujos, pero la invención no está limitada por los mismos, sino solamente por las reivindicaciones. Los dibujos descritos son sólo esquemáticos y no son limitativos.

Además, los términos primero, segundo, tercero y similares, en la descripción y en las reivindicaciones, son usados para distinguir entre elementos similares y no necesariamente para describir un orden secuencial o cronológico. Debe entenderse que los términos usados son intercambiables bajo circunstancias apropiadas y que las realizaciones de la invención descrita en la presente memoria son capaces de funcionar en otras secuencias que las descritas e ilustradas en la presente memoria.

Cabe destacar que la expresión "que comprende", usada en las reivindicaciones, no debería interpretarse como estando restringida a los medios indicados más adelante; no excluye otros elementos o etapas. De esta manera, debe interpretarse que especifica la presencia de las características, enteros, etapas o componentes indicados, tal como se refiere, pero no excluye la presencia o adición de una o más características, enteros, etapas o componentes diferentes, o grupos de los mismos. De esta manera, el alcance de la expresión "un dispositivo que comprende medios A y B" no debería limitarse a dispositivos que consisten sólo en componentes A y B. Significa que, con respecto a la presente invención, los únicos componentes relevantes del dispositivo son A y B.

Ahora, la invención será descrita por medio de una descripción detallada de diversas realizaciones de la invención. Es evidente que otras realizaciones de la invención pueden ser configuradas según el conocimiento de las personas con conocimientos en la materia, sin apartarse de las enseñanzas técnicas de la invención, estando limitada la invención sólo por los términos de las reivindicaciones adjuntas.

En un aspecto, la presente invención propone una codificación y decodificación de audio multicanal.

1. Principio básico.

El presente aspecto de la presente invención se refiere a la codificación de una pluralidad de canales en una señal de audio mezclada de manera que los flujos de audio originales pueden ser reconstruidos a partir de la señal de audio mezclada, y la decodificación de al menos uno de entre la pluralidad de canales a partir de la señal de audio mezclada.

- 5 Un ejemplo de la presente invención se describe con referencia a la codificación y decodificación de 4 canales, pero la presente invención no está limitada en este sentido. Cuantos más canales sean grabados, más muestras tienen que ser calculadas usando un filtro y menor puede ser la calidad de la música reproducida, pero, en algunas aplicaciones, la calidad de la música no es de importancia primordial. El formato de codificación de la presente invención está basado en la mezcla de 2 o más flujos digitales de audio, de tal manera que los flujos de audio originales pueden ser reconstruidos a partir de la señal de audio mezclada.

Como un ejemplo para ilustrar esto, se parte de 2 flujos A y B de audio digital, mono, de 44.1 KHz, de 16 bits.

Al mezclar dos señales de audio, se usa una técnica de aproximación.

Se usa Modulación por codificación de impulsos (PCM) para muestrear audio analógico a un código digital, típicamente, a 8.000 muestras/segundo. La cadena de audio digital es denominada, entonces, un flujo PCM.

- 15 En la Figura 4, un primer flujo A de audio PCM se muestra en el gráfico, como una línea gris oscura (serie 1). Las muestras del primer flujo A de audio son: $A_0, A_1, A_2, A_3, A_4, A_5, \dots$ A partir de esta serie de muestras, se genera un nuevo flujo A' (gris claro, serie 2). El nuevo flujo A' es una aproximación del primer flujo A de audio, usando una interpolación lineal para aproximar las muestras impares del flujo A' recién generado, mientras se copian las muestras pares. Las muestras del flujo A' son: $A'_0, A'_1, A'_2, \dots$ con $A'_{2i} = A_{2i}$ y $A'_{2i+1} = (A'_{2i} + A'_{2i+2})/2$. De esta manera, todas las muestras pares de la serie de muestras recién generada son idénticas a los datos originales y todas las muestras impares se definen como la interpolación lineal de la muestra anterior y la muestra siguiente.

- 20 En la Figura 5, un segundo flujo B de audio PCM se muestra en el gráfico, como una línea gris oscura (serie 1). Las muestras del segundo flujo B de audio son: $B_0, B_1, B_2, B_3, B_4, B_5, \dots$ A partir de esta serie de muestras, se genera un nuevo flujo B' (gris claro, serie 2). El nuevo flujo B' es una aproximación del segundo flujo B de audio, usando una interpolación lineal para aproximar las muestras pares del flujo B' recién generado, mientras se copian las muestras impares. Las muestras del flujo B' son: $B'_0, B'_1, B'_2, \dots$ con $B'_{2i+1} = B_{2i+1}$ y $B'_{2i} = (B'_{2i-1} + B'_{2i+1})/2$. De esta manera, todas las muestras impares son idénticas a los datos originales y todas las muestras pares se definen como la interpolación lineal de la muestra anterior y la muestra siguiente.

- 25 En la Figura 6, ambos flujos A y B originales se mezclan (sumados) para obtener un primer flujo A+B mezclado (gris oscuro, serie 1). Los flujos A' y B' aproximados son mezclados (sumados) para obtener un segundo flujo A'+ B' mezclado (gris claro, serie 2). El segundo flujo A'+ B' mezclado puede ser diferente del primer flujo A+B mezclado para cada muestra, ya que A' o B' pueden diferir de las muestras originales, pero, en general, el segundo flujo mezclado es todavía una buena aproximación del flujo A + B mezclado original (gris oscuro, serie 1).

- 30 Para des-mezclar el flujo A'+ B' mezclado, a fin de obtener las señales A', B' mono, aproximadas, a partir del flujo A' + B' mezclado, se necesitan la primera muestra original $A_0 = A'_0$ del primer flujo A de audio y la segunda muestra $B_1 = B'_1$ del segundo flujo B de audio.

- 35 El des-mezclado de las señales (mono) a partir de A'+ B' puede ser realizado de la manera siguiente: Las muestras del segundo flujo A'+ B' mezclado son: $A'_0 + B'_0, A'_1 + B'_1, A'_2 + B'_2, A'_3 + B'_3, A'_4 + B'_4, A'_5 + B'_5, \dots$ Cuando se tiene una copia de la primera muestra $A_0 = A'_0$ del primer flujo A de audio y de la segunda muestra $B_1 = B'_1$ del segundo flujo B de audio, los flujos A' y B' pueden ser reconstruidos.

- 40 1. con $(A'_0 + B'_0)$ y (A'_0) conocidos, puede obtenerse B'_0
 2. con $(A'_1 + B'_1)$ y (B'_1) conocidos, puede obtenerse (A'_1)
 3. con (A'_0) y (A'_1) conocidos, puede obtenerse (A'_2) , usando $A'_1 = (A'_0 + A'_2)/2$ o, de esta manera, $A'_2 = (2 A'_1 - A'_0)$
 4. con $(A'_2 + B'_2)$ y (A'_2) conocidos, puede obtenerse (B'_2)
 45 5. con (B'_1) y (B'_2) conocidos, puede obtenerse (B'_3) , usando $B'_2 = (B'_1 + B'_3)/2$ o, de esta manera, $B'_3 = (2 B'_2 - B'_1)$
 6. con $(A'_3 + B'_3)$ y (B'_3) conocidos, puede obtenerse A'_3
 7. etc ..

Si este procedimiento de mezclado de un flujo A, B de audio original con un flujo A', B' de audio interpolado es repetido para otro par de flujos (C, D) PCM (mono), se obtiene un segundo flujo C' + D' PCM (mono) mezclado. Estos 2 flujos A' + B', C' + D', PCM (mono) mezclados constituyen un flujo PCM estéreo equivalente, que son los datos de audio PCM básicos generados por el codificador CD+G(w) y usados por las herramientas de creación CD+G.

- 5 Con este principio básico en mente, puede concluirse que al menos algunos duplicados ($A_0 = A'_0$ y $B_1 = B'_1$) (idénticos) de las primeras muestras PCM iniciales a partir de los flujos originales son requeridos como parte de los parámetros de filtro, con el fin de poder decodificar o des-mezclar los flujos de audio PCM mezclados.

2. Restricciones y limitaciones del principio básico

- 10 La idea básica, tal como se ha explicado en la sección anterior, para mezclar/des-mezclar 2 flujos PCM mono, tiene diversas limitaciones y restricciones.

- 15 La aproximación de un flujo PCM, usando una interpolación lineal para aproximar las muestras impares de un flujo (mientras se copian las muestras pares), y viceversa para el otro flujo PCM, resulta en una pérdida de información de frecuencia más elevada del flujo PCM original. Por lo tanto, puede usarse una técnica de aproximación más compleja, que incluye la optimización de los parámetros del filtro usando criterios tales como la minimización de los errores de interpolación, o la optimización de las características de frecuencia, que pueden incluir características perceptuales como parte de los criterios de optimización.

- 20 El mezclado de 2 flujos PCM puede resultar en rebasamientos. Cada uno de los flujos PCM del presente ejemplo es un flujo de 16 bits, de 44,1 kHz, con valores enteros de 16 bits, con signo, con un máximo de 32.767 y un mínimo de -32.768. La suma de 2 flujos PCM, donde cada flujo podría tener su máximo de 32.767 en el mismo momento, proporciona un valor total de 65.534, que excede el máximo PCM de 16 bits (con signo). Para evitar dichos rebasamientos, puede usarse un esquema de compresión dinámica, aplicado a los datos PCM, previamente a mezclar todos estos flujos en un flujo estéreo convertido desde multicanal.

- 25 Los parámetros del filtro usados por técnicas más complejas en la aproximación del flujo PCM, son el resultado de una optimización y se representan como valores en coma flotante. Sin embargo, para hacer que el procesamiento de codificación/decodificación sea independiente de implementaciones aritméticas en coma flotante, la codificación CD+G(w) usada con la presente invención se basará en cálculos enteros y, de esta manera, los parámetros tienen que ser convertidos o representados por un número entero, tanto para la codificación como para la decodificación.

- 30 Finalmente, el des-mezclado de las muestras PCM con este algoritmo básico puede resultar en errores, cuando se produce un error durante la lectura desde el disco CD+G. La naturaleza de este procedimiento es tal que todas las muestras reconstruidas después de un único error de lectura, serán también incorrectas, ya que las muestras PCM siguientes son reconstruidas usando la muestra anterior con su error, dando lugar a errores adicionales.

La técnica de codificación/decodificación CD+G(w) avanzada según la presente invención abordará todas estas deficiencias y/o limitaciones.

3. Codificación y decodificación avanzada según las realizaciones de la presente invención

- 35 a. Parámetros del filtro

Para el primer flujo PCM, las muestras impares deben ser interpoladas. La interpolación lineal, tal como se explica en la sección acerca del principio básico, es remplazada por un filtro FIR, donde los parámetros del filtro son optimizados tal como se muestra en la Figura 7.

- 40 Los parámetros del filtro son usados para aproximar las muestras impares, por ejemplo, la muestra en A_5 es generada a partir de las muestras pares A_0 , A_2 , A_4 y A_6 . $A'_5 = \alpha \cdot A_6 + \beta \cdot A_4 + \gamma \cdot A_2 + \delta \cdot A_0$.

La fórmula general para la aproximación de una muestra impar A'_{2i+5} es:

$$A'_{2i+5} = \alpha \cdot A_{2i+6} + \beta \cdot A_{2i+4} + \gamma \cdot A_{2i+2} + \delta \cdot A_{2i} \text{ o } A'_{2i+5} = (1/\alpha) \cdot (A'_{2i+5} - \beta \cdot A_{2i+4} - \gamma \cdot A_{2i+2} - \delta \cdot A_{2i})$$

- 45 Esta fórmula puede ser usada para el algoritmo básico cuando (α , β , γ , δ) se definen como (0,5, 0,5, 0, 0), pero con el codificador según las realizaciones de la presente invención, (α , β , γ , δ) es el resultado de una optimización, que se explicará más adelante. Además, debido a que el decodificador necesitará ($1/\alpha$), se definen restricciones sobre los conjuntos de parámetros, para evitar valores de α demasiado cercanos a 0, para eliminar divisiones por cero. Además, debido a que se usan cálculos enteros en el esquema de codificación y decodificación final, deben establecerse restricciones todavía más altas sobre los valores α para evitar cálculos que introducen errores de redondeo demasiado altos.

La fórmula para la aproximación de una muestra par B'_{2i+6} es:

$$B'_{2i+6} = \alpha \cdot B_{2i+7} + \beta \cdot B_{2i+5} + \gamma \cdot B_{2i+3} + \delta \cdot B_{2i+1} \text{ o } B'_{2i+7} = (1/\alpha) \cdot (B'_{2i+6} - \beta \cdot B_{2i+5} - \gamma \cdot B_{2i+3} - \delta \cdot B_{2i+1})$$

Un primer flujo A PCM, con muestras $A_0, A_1, A_2, A_3, A_4, A_5, A_6, A_7, \dots$, es usado para generar un nuevo flujo A' . El nuevo flujo A' es una aproximación del primer flujo A, en el que las muestras pares se copian desde el primer flujo A PCM, y las muestras impares son aproximadas. Las muestras de A' son: $A'_0, A'_1, A'_2, A'_3, A'_4, A'_5, A'_6, A'_7, \dots$ con $A'_{2i} = A_{2i}$ y A'_{2i+5} se define por la fórmula anterior para $i \geq 0$. $A'_i = A_i$ para $i < 5$. Un segundo flujo B PCM, con muestras $B_0, B_1, B_2, B_3, B_4, B_5, B_6, B_7, \dots$ es usado para generar un nuevo flujo B' . El nuevo flujo B' es una aproximación del segundo flujo B, en el que las muestras impares se copian desde el segundo flujo B PCM, y las muestras pares son aproximadas. Las muestras de B' son: $B'_0, B'_1, B'_2, B'_3, B'_4, B'_5, B'_6, B'_7, \dots$ con $B'_{2i+1} = B_{2i+1}$ y B'_{2i+6} definidos por una fórmula equivalente para todo $i \geq 0$. $B'_i = B_i$ para $i < 6$.

b. Operaciones de des-mezclado

Al mezclar estos flujos A' y B' recién generados, se obtiene: $A'_0 + B'_0, A'_1 + B'_1, A'_2 + B'_2, A'_3 + B'_3, A'_4 + B'_4, A'_5 + B'_5, A'_6 + B'_6, A'_7 + B'_7, A'_8 + B'_8, A'_9 + B'_9, \dots$ Usando una copia de las primeras muestras pares A_0, A_2, A_4 del primer flujo A de audio, y una copia de las primeras muestras impares B_1, B_3, B_5 , del segundo flujo B de audio, y los conjuntos de parámetros del filtro ($\alpha_A, \beta_A, \gamma_A, \delta_A$) y ($\alpha_B, \beta_B, \gamma_B, \delta_B$), puede definirse el siguiente esquema de des-mezclado:

1. con $(A'_0 + B'_0)$ y (A'_0) conocidos, puede obtenerse B'_0
2. con $(A'_1 + B'_1)$ y (B'_1) conocidos, puede obtenerse A'_1
3. con $(A'_2 + B'_2)$ y (A'_2) conocidos, puede obtenerse B'_2
4. con $(A'_3 + B'_3)$ y (B'_3) conocidos, puede obtenerse A'_3
5. con $(A'_4 + B'_4)$ y (A'_4) conocidos, puede obtenerse B'_4
6. con $(A'_5 + B'_5)$ y (B'_5) conocidos, puede obtenerse A'_5
7. con A'_0, A'_2, A'_4, A'_5 conocidos, puede calcularse A'_6 usando $A'_6 = (1/\alpha_A) \cdot (A'_5 - \beta_A \cdot A'_4 - \gamma_A \cdot A'_2 - \delta_A \cdot A_0)$
8. con $(A'_6 + B'_6)$ y (A'_6) conocidos, puede obtenerse B'_6
9. con B'_1, B'_3, B'_5, B'_6 conocidos, puede calcularse B'_7 usando $B'_7 = (1/\alpha_B) \cdot (B'_6 - \beta_B \cdot B'_5 - \gamma_B \cdot B'_3 - \delta_B \cdot B'_1)$
10. con $(A'_7 + B'_7)$ y (B'_7) conocidos, puede obtenerse A'_7
11. con A'_2, A'_4, A'_6, A'_7 conocidos, puede calcularse A'_8 usando $A'_8 = (1/\alpha_A) \cdot (A'_7 - \beta_A \cdot A'_6 - \gamma_A \cdot A'_4 - \delta_A \cdot A'_2)$
12. con $(A'_8 + B'_8)$ y (A'_8) conocidos, puede obtenerse B'_8
13. con B'_3, B'_5, B'_7, B'_8 conocidos, puede calcularse B'_9 usando $B'_9 = (1/\alpha_B) \cdot (B'_8 - \beta_B \cdot B'_7 - \gamma_B \cdot B'_5 - \delta_B \cdot B'_3)$
14. con $(A'_9 + B'_9)$ y (B'_9) conocidos, puede obtenerse A'_9
15. etc ..

En esta etapa, se puede concluir que junto con los conjuntos de parámetros ($\alpha_A, \beta_A, \gamma_A, \delta_A$) y ($\alpha_B, \beta_B, \gamma_B, \delta_B$), se necesitan también una copia de 3 muestras de la primera serie (A'_0, A'_2, A'_4) y 3 muestras de la segunda serie (B'_1, B'_3, B'_5).

c. Secciones de tiempo y asignación de parámetro entero

Una de las restricciones del principio básico era que no se recupera de un error. Una forma de limitar los efectos de un error de lectura (leyendo las muestras digitalmente desde el CD) es limitar el uso de estos parámetros ($\alpha_A, \beta_A, \gamma_A, \delta_A$) y ($\alpha_B, \beta_B, \gamma_B, \delta_B$) y la series de muestras iniciales (A'_0, A'_2, A'_4) y (B'_1, B'_3, B'_5) a un número restringido de muestras PCM. Se ha encontrado que 40 ms de muestras de audio (3x 588 muestras) es una longitud aceptable para una serie de muestras. Dicho un grupo de muestras se denominará una "sección". Esto permite reducir el efecto de un error a una única sección, (máx 40 ms), así como optimizar un grupo de parámetros de filtro para una cantidad menor de muestras, proporcionando un mejor comportamiento, por ejemplo, en términos de errores de interpolación o características de frecuencia. Para cada una de estas secciones, se necesitarán las primeras 3 muestras, así como los parámetros del filtro.

Tal como se ha indicado anteriormente, los parámetros del filtro ($\alpha, \beta, \gamma, \delta$) están optimizados y resultan en números en coma flotante. Estos números son representados por números enteros, ya que el algoritmo de codificación y

decodificación está implementado con operaciones de enteros, para hacerlo portable e independiente de las implementaciones en coma flotante. Se usa la siguiente conversión para representar estos parámetros usando un número de 16 bits, con signo, y un valor de 2 bits (número de la base). Por ejemplo, α es convertido en un valor entero A de (-32.768) a 32.767 mientras que A_b recibe los valores 0, 1, 2 ó 3. El procedimiento se explica a continuación:

5 En primer lugar, se define el valor de la base Ab:

Quando $\alpha > 0$: $Ab = (\text{unsigned char}) (\text{fabs}(\alpha + 0,99999999999999999999)) - 1;$

Si no, si $\alpha < 0$: $Ab = (\text{unsigned char}) (\text{fabs}(\alpha - 0,99999999999999999999)) - 1$;

Si no: $Ab = 0$;

Nota: fabs(x) devuelve el valor absoluto del valor x en coma flotante.

10 A continuación,

cuando $Ab > 7$: condición de rebasamiento, la optimización de parámetros no debería recibir esos valores

cuando $Ab = 7, 6, 5$ ó 4 , Ab se restablece a $Ab = 3$

cuando $Ab = 3$ ó 2 , Ab se restablece a $Ab = 2$

cuando $Ab = 1$, Ab no cambia.

15 cuando $Ab = 0$, Ab no cambia.

O para resumir

$$Ab = 0 \text{ cuando } 1 \geq \alpha \geq (-1),$$

$Ab = 1$ cuando $2 \geq \alpha > 1$ ó $(-1) > \alpha \geq (-2)$.

$Ab = 2$ cuando $4 \geq \alpha > 2$ ó $(-2) > \alpha \geq (-4)$,

20 $Ab = 3$ cuando $8 > \alpha > 4$ ó $(-4) > \alpha > (-8)$.

Finalmente, α es convertido en un entero de 16 bits, con signo, usando la fórmula siguiente:

$$A = (\text{short}) (\alpha \times (32.768 / (1 \ll Ab))).$$

De esta manera, α es convertido en un par (A, Ab), β es convertido en un par (B, Bb), γ es convertido en un par (C, Cb) y δ es convertido en un par (D, Db). Los números de base Ab, Bb, Cb, Db son fusionados en una BASE = Ab | (Bb << 4) | (Cb << 8) | (Db << 12) de 16 bits.

25 (Cb << 8) | (Db << 12) de 16 bits.

La conclusión en esta etapa es la siguiente:

Para cada sección de 40 ms, a partir de un único flujo PCM, mono, de 16 bits, de 44,1 KHz, o 3x 588 muestras (mono) de 16 bits, el codificador según las realizaciones de la presente invención define un conjunto de parámetros de filtro (α , β , γ , δ) y convierte estos parámetros en cuatro números (A, B, C, D) de 16bit, con signo, y un número BASE de 16 bits. Además, para la operación de des-mezclado, se requieren 3 muestras de 16 bits; serie (A'₀, A'₂, A'₄). Esto suma un total de 8 valores (4 + 1 + 3) de 16 bits, necesarios para las operaciones de des-mezclado. Para el segundo flujo PCM de 16 bits, mono, de 44,1 KHz, un conjunto similar de 8 valores de 16 bits son generados por el codificador según las realizaciones de la presente invención.

30 Además, para la operación de des-mezclado, se requieren 3 muestras de 16 bits; serie (A₀, A₂, A₄). Esto suma un total de 8 valores (4 + 1 + 3) de 16 bits, necesarios para las operaciones de des-mezclado. Para el segundo flujo PCM de 16 bits, mono, de 44,1 KHz, un conjunto similar de 8 valores de 16 bits son generados por el codificador según las realizaciones de la presente invención.

En total, el codificador según las realizaciones de la presente invención es capaz de mezclar una pluralidad, por ejemplo 4, flujos PCM, mono, de 16 bits en un flujo PCM, estéreo, de 16 bits. Un segundo de música es dividido en secciones de 40 ms (o 25 secciones por segundo). Cada sección requiere 4x 8 x 16 bits valores/parámetros.

35 flujos PCM, mono, de 16 bits en un flujo PCM, estéreo, de 16 bits. Un segundo de música es dividido en secciones de 40 ms (o 25 secciones por segundo). Cada sección requiere $4 \times 8 \times 16$ bits valores/parámetros.

En esta etapa, el codificador requiere 25x 4x 8x 2x bytes de datos adicionales por segundo, ó 1.600 bytes por segundo.

d. Esquemas avanzados de mezclado y otros parámetros según las realizaciones de la presente invención

Hasta ahora, se han proporcionado ejemplos de mezclado de dos flujos PCM, mono, de 16 bits, en otro flujo PCM, mono, de 16bits, pero el codificador según las realizaciones de la presente invención soporta diferentes esquemas de mezclado y des-mezclado o "Modos de mezclado". El Modo de mezclado seleccionado es parte de los parámetros CD+G(w), usando un parámetro de Modo de mezclado de 4 bits.

de 16bits, pero el codificador según las realizaciones de la presente invención soporta diferentes esquemas de mezclado y des-mezclado o “Modos de mezclado”. El Modo de mezclado seleccionado es parte de los parámetros CD+G(w), usando un parámetro de Modo de mezclado de 4 bits.

Los Modos de mezclado siguientes se proporcionan como ejemplos de referencia y pueden ser usados con un codificador según la presente invención;

Modo de mezclado 0: **Solo estéreo (Instrumento/Voz), mezclado con fondo en estéreo.**

- Instrumento o voz solista en estéreo, los canales izquierdo y derecho son (Solo-L, Solo-R)

5 - Fondo en estéreo, los canales izquierdo y derecho (Bckg-L, Bckg-R)

Se mezclan en: Canales mezclados en estéreo (Solo-L + Bckg-L, Solo-R + Bckg-R).

No se requieren operaciones adicionales de mezclado o des-mezclado.

Modo de mezclado 1: **Voz e instrumento mono mezclados con fondo en estéreo.**

- Canal de voz mono es (Voc)

10 - Canal de instrumento mono es (Instr)

- Fondo en estéreo, los canales izquierdo y derecho son (Bckg-L, Bckg-R)

Se mezclan en: Canales estéreos mezclados (Bckg-L + a.Voc + b.Instr, Bckg-R + c.Voc + d.Instr)

La panoramización de instrumento y voz depende de los valores para (a, b, c, d), donde a, b, c y d son tales que no hay ningún número t que cumpla $a = t.c$ y $b = t.d$. Los valores típicos son:

15 (a, b, c, d) = (1, 1, 1, 2): panoramiza la voz en el centro, el instrumento más a la derecha.

(a, b, c, d) = (1, 2, 1, 1): panoramiza la voz en el centro, el instrumento más a la izquierda.

(a, b, c, d) = (2, 1, 1, 2): panoramiza la voz más a la izquierda, el instrumento más a la derecha.

(a, b, c, d) = (1, 2, 2, 1): panoramiza la voz más a la derecha, el instrumento más a la izquierda.

a, b, c ó d pueden ser > 2, pero estos ejemplos cubrirán la mayoría de los casos.

20 Este esquema de mezclado requiere algunas operaciones adicionales, a fin de preparar los flujos originales a ser convertidos desde multicanal. En primer lugar, se generan 2 nuevos flujos mono;

- LeftVI = a.Voc + b.Instr

- RightVI = c.Voc + d.Instr

25 Estos nuevos flujos se mezclaron con los canales de fondo (Bckg-L, Bckg-R) en (Bckg-L + a.Voc + b.Instr, Bckg-R + c.Voc + d.Instr). El des-mezclado de estos flujos suministrará LeftVI y RightVI, que pueden ser usados para reconstruir las pistas vocales e instrumentales, usando esta fórmula:

- Instr = $(a / (d.a - c.b)) \cdot (RightVI - (c/a) \cdot LeftVI)$

- Voc = $(b / (c.b - a.d)) \cdot (RightVI - (d/b) \cdot LeftVI)$

Modo de mezclado 2: **4 canales mono independientes mezclados a un canal estéreo convertido desde multicanal**

30 - Canal mono 1 es (C1)

- Canal mono 2 es (C2)

- Canal mono 3 es (C3)

- Canal mono 4 es (C4)

Se mezclan a: Canales en estéreo (a.C1 + b.C2 + a'.C3 + b'.C4, c.C1 + d.C2 + c'.C3 + d'.C4)

35 Este esquema de mezclado requiere algunas operaciones adicionales con el fin de preparar los flujos originales a convertir a estéreo. En primer lugar, se generan 4 nuevos flujos mono;

- Left1 = a.C1 + b.C2

$$\text{- Right1} = c.C1 + d.C2$$

$$\text{- Left2} = a'.C3 + b'.C4$$

$$\text{- Right2} = c'.C3 + d'.C4$$

5 Estos canales se mezclan a (Left1 + Left2, Right1 + Right2). El par (a, c) define la panoramización del primer canal C1 al estéreo convertido desde multicanal. (b, d), (a'c') y (b'd'), respectivamente, definen la panoramización para C2, C3 y C4. Una vez más, se aplica la restricción de que no hay números t & t' que cumplan $a = t.c$ y $b = t.d$ y $a' = t'.c'$ y $b' = t'.d'$. Los valores típicos son:

(a, c) = (3, 1): panoramiza el Canal 1 a la izquierda

(b, d) = (1, 3): panoramiza el Canal 2 a la derecha

10 (a', c') = (2, 1): panoramiza el Canal 3 al centro-izquierda

(a', c') = (1, 2): panoramiza el Canal 4 al centro-derecha

A partir de la mezcla en estéreo (Left1 + Left2, Right1 + Right2), Left1 y Left2, Right 1 y Right 2 pueden ser des-mezclados. A partir de estos canales des-mezclados, C1, C2, C3 y C4 pueden ser reconstruidos usando:

$$C1 = (b / (c.b - a.d)).(Right1 - (d/b).Left1)$$

$$15 \quad C2 = (a / (d.a - c.b)).(Right1 - (c/a).Left1)$$

$$C3 = (b' / (c'.b' - a'.d')).(Right2 - (d'/b').Left2)$$

$$C4 = (a' / (d'.a' - c'.b')).(Right2 - (c'/a').Left2)$$

Modo de mezclado 3: Audio cuadrafónico: canales estéreo frontal y estéreo posterior

- Los canales estéreos frontales izquierdo y derecho son (Fr-L, Fr-R)

20 - Los canales estéreos posteriores izquierdo y derecho son (Rr-L, Rr-R)

Se mezclan a: (Fr-L + a.Rr-L, Fr-R + a.Rr-R)

Típicamente, los canales posteriores son atenuados antes del mezclado a una grabación en estéreo: (Fr-L + a.Rr-L, Fr-R + a.Rr-R). El parámetro "a" indica una atenuación de 0 dB a -255 dB

25 Modos de mezclado 4-15: Otros modos de mezclado son libres para ser definidos para aplicaciones dedicadas, y los codificadores/decodificadores estándar pueden ser compatibles o no con estos modos extra.

Debido a que el codificador según las realizaciones de la presente invención es capaz de mezclar y des-mezclar diferentes esquemas de panoramización, el decodificador requiere parámetros adicionales según las realizaciones de la presente invención y son generados por el codificador para seleccionar el esquema de des-mezclado correcto. Las opciones adicionales podrían consistir en incluir información de panoramización envolvente dinámica usada por el decodificador para posicionar, de manera dinámica, los canales independientes dentro del espacio auditivo 3D. Por ejemplo, usando 1 byte para la panoramización izquierda/derecha dinámica, y otro byte para la panoramización frontal/posterior dinámica para un canal específico. Otra opción consiste en almacenar los parámetros adicionales del filtro, por ejemplo, para las características del filtro que pueden ser útiles cuando se aplican operaciones de cambio de tono o expansión temporal a los canales de audio des-mezclados. Las técnicas de procesamiento de audio, tales como el cambio de tono, son útiles para el karaoke, y la expansión temporal es útil cuando el CD+G(w) es usado como base para la formación musical para cantar y tocar en grupo. Los efectos de procesamiento crean, de alguna manera, distorsión, y filtros adicionales pueden ser útiles para reducir los artefactos de estos efectos. El uso de dichos conjuntos de parámetros adicionales sólo se incluye en la presente memoria como opciones extensionales.

40 Para poder vincular los grupos de parámetros al número de sección de audio correcto, el número de sección puede estar incluido como parte de los parámetros. Estos números de sección están relacionados con el inicio de esa pista específica. (La primera sección de una pista en estéreo es el número 0). Para estos números de sección, se asignan 18 bits (el máximo es 131.071). (80 mins = 80x60x25 = 120.000 secciones). A continuación, el formato CD+G(w) puede incorporar también datos MIDI y letras. Si estos datos están presentes, se indica usando un único bit para MIDI, y otro bit para las letras.

45 En resumen, para cada sección de audio, se usan 8 bytes adicionales para especificar estos parámetros:

byte1: ((Número de sección & 0x00FF))

byte2: ((Número de sección & 0xFF00) >> 8)

byte3: ((Número de sección & 0x3000) >> 16) | (Modo-mezclado << 2) | (MIDI << 6) | (Letras << 7)

Los bytes siguientes dependen del modo de mezclado:

5 Modo de mezclado 0:

byte4 - byte8: podrían ser usados para almacenar los parámetros de expansión temporal o filtros de cambio de tono

Modo de mezclado 1:

byte4: (a << 6 | b << 4 | c << 2 | d) (Esquema de mezclado a.V + b.I y c.V + d.I)

10 byte5 - byte8: usados para los filtros de desplazamiento temporal/expansión temporal o panoramización dinámica de audio de 1 ó 2 canales, 1 byte para Left/Right: (0x80: centro, 0xFF: derecha, 0x00: izquierda), otro byte para Parte frontal/Parte posterior: (0x80: centro, 0xFF: frontal, 0x00: posterior)

Modo de mezclado 2:

byte4: (a << 6 | b << 4 | c << 2 | d) (Esquema de mezclado a.C1 + b.C2 y c.C1 + d.C2)

byte5: (a' << 6 | b' << 4 | c' << 2 | d') (Esquema de mezclado a'.C3 + b'.C4 y C'.C3 + D'.C4)

15 byte6 - byte8: usados para los filtros desplazamiento temporal/expansión temporal o panoramización dinámica de 1 ó 2 canales, 1 byte para Left/Right: (0x80: centro, 0xFF: derecha, 0x00: izquierda), otro byte para Parte frontal/Parte posterior: (0x80: centro, 0xFF: frontal, 0x00: posterior)

Modo de mezclado 3:

byte4: atenuación de sonido envolvente de canal de audio (0 dB a -255 dB)

20 byte5 - byte8: disponible

Modo de mezclado 4-15: modos a definir según la aplicación.

byte4-byte8: a definir

25 Estos 8 bytes adicionales por sección, 25 secciones por segundo, suman otros 200 bytes por segundo, de manera que el total de datos de parámetros según las realizaciones de la presente invención, incluyendo los parámetros de filtrado de audio, suman $1.600 + 200 = 1.800$ bytes por segundo.

e. Compresores

El mezclado de los flujos PCM puede introducir rebasamientos. Un número PCM, de 16 bits, con signo, está limitado a valores entre (-32.768) y (32.767). Con el fin de explicar los esquemas de compresión, los flujos PCM usados en el ejemplo de mezclado siguiente se definen como se indica a continuación:

30 Un flujo PCM (voz), mono, de 16 bits, y un flujo PCM (instrumento), mono, de 16 bits son mezclados con un flujo PCM (audio de fondo-L/fondo-R), de 16 bits.

35 La primera pasada del codificador según las realizaciones de la presente invención es para definir los parámetros del filtro para Voz, Instrumento, Fondo-L y Fondo-R. Los flujos son convertidos primero a datos en coma flotante, y se asignan a una escala de valores entre (-1) y (1). Se usa la aproximación de muestra "impar" para los flujos de voz e instrumento, y la aproximación de muestra "par" para el Fondo-L y Fondo-R. Estos flujos aproximados recién generados son Voc', Instr', BckgL' y BckgR'. Debido a la naturaleza de la aproximación, pueden haberse introducido rebasamientos durante la generación de estos nuevos flujos, obteniendo valores > 1 ó $< (-1)$. Por lo tanto, los nuevos datos son escalados, típicamente mediante la multiplicación por ($\frac{1}{2}$) para volver a un intervalo de valores entre (-1) y (1) y, finalmente, los flujos son convertidos en flujos PCM, de 24 bits, con signo, con valores entre (-8.388.608) y (8.388.607) para añadir resolución.

40 A continuación, las marcas de tiempo de estas muestras que generan rebasamientos son identificadas antes de mezclarlas en la conversión final en estéreo desde multicanal. Para cada uno de estos flujos, de manera individual, se definen los factores de compresión asociados con las marcas de tiempo requeridas para reducir los rebasamientos. Cada muestra que no genera un rebasamiento recibirá un factor de escala de 1, pero las muestras que generan rebasamientos

reciben un factor de escala < 1 . Para cada flujo, se define un valor umbral máximo. Todas las muestras en las que, por ejemplo, $|Voc_i| > UmbralMax$, recibe un factor de escala definido por $(UmbralMax / |Voc_i|)$. Los flujos comprimidos se definen como Voc", Instr", BckgL" y BckgR".

5 La conversión final desde multicanal al flujo en estéreo es otra fuente de generación de rebasamientos, que requiere una compresión adicional. Por ejemplo, si se toma un esquema de mezclado como el siguiente: $(BckgL'' + Voc'' + 2xInstr'')$ y $(BckgR'' + Voc'' + Instr'')$. Cuando $|(BckgL''_i + Voc''_i + 2xInstr''_i)| > UmbralMax$ o $|(BckgR''_i + Voc''_i + Instr''_i)| > UmbralMax$, se define un factor de escala adicional. En el caso en el que ambas muestras mezcladas (izquierda y derecha) requieren compresión, se selecciona la compresión más alta y se aplica a ambas. En el caso en el que sólo una muestra mezclada requiere compresión, esa compresión es aplicada aún a ambas muestras mezcladas (ambas muestras izquierda y derecha). Como resultado de la compresión, se obtienen nuevos flujos, Voc*, Instr*, BckgR* y BckgL*, donde, por ejemplo, $Voc^*_i = Voc_i \times Vcompr_i$ y Vcompr es el resultado de multiplicar todos los factores de escala aplicados a esa muestra particular de ese flujo. La secuencia de todos estos factores de compresión para generar, por ejemplo, Voc*, como tal, constituyen un nuevo flujo Vcompr, en el que la mayoría de los factores de escala serán $= 1$ y aquellos que requieren compresión tendrán un factor de escala < 1 .

15 En esta etapa, se obtienen 4 flujos de datos de factores de compresión, Vcompr, lcompr, BRcompr y BLcompr, que definen los factores de escala requeridos para cada muestra en cada flujo de audio PCM. Estos son en realidad una compresión de "impulsos", ya que funcionarán sobre las muestras específicas que introducen rebasamientos en el mezclado. Al aplicar dicha compresión sólo a muestras "aisladas" se introducirá una distorsión sustancial o artefactos en la grabación de audio, ya que una compresión de "impulsos" de la muestra no sería diferente de una discontinuidad del flujo de audio. Por lo tanto, debe aplicarse un filtro de efecto temporal, que extiende el efecto de compresión a lo largo de un periodo de tiempo típico de 100 milisegundos antes y después de estas marcas de tiempo específicas de estas muestras.

La Figura 8 proporciona una visión general de los factores de compresión, donde el efecto de compresión se aplica sobre varias muestras consecutivas.

25 En la sección superior de la Figura 8, la línea con rombos negros (serie 1) representa las muestras del flujo original. La línea con cuadrados grises (serie 2) es el resultado de una (mala) aproximación. Si se define un valor umbral de 60, hay presentes 2 muestras en las que se necesita la compresión, en la posición 4 de las muestras originales y en la posición 10, debido a la aproximación.

30 En la sección inferior de la Figura 8, la línea con rombos negros (serie 1) representa los factores de compresión de "impulsos", mientras que la línea gris claro (serie 2) define a una mejor compresión que se usará según una realización de la presente invención. Se aplica una compresión gradual sobre un número de muestras antes y después de los casos de compresión de "impulsos", por ejemplo, sobre las dos muestras antes y sobre dos muestras después de la muestra en la que se necesita compresión.

f. Optimización de parámetros.

35 Tal como se ha explicado en la sección anterior acerca de los parámetros del filtro usados por el codificador/decodificador según las realizaciones de la presente invención, los parámetros del filtro (α , β , γ , δ) tienen que ser optimizados. Los parámetros son usados para la aproximación de una sección de 1 flujo PCM mono, de 16bits, de 44,1 kHz. Dicha una sección es equivalente a 3×588 muestras o 40 ms. La optimización de los parámetros requiere criterios de optimización. El apartado siguiente proporciona una visión general de los diferentes criterios útiles para la optimización.

40 A modo de referencia, la sección original de ese flujo PCM mono se define como A, con muestras A_i y $0 \leq i < 1.764$, y los parámetros del filtro (α , β , γ , δ). El resultado de la aproximación es un nuevo flujo, A'_i ($0 \leq i < 1.764$), con $A'_{2i} = A_{2i}$, $A'_{2i+5} = \alpha.A_{2i+6} + \beta.A_{2i+4} + \gamma.A_{2i+2} + \delta.A_{2i}$ y $A'_i = A_i$ para $i \leq 5$.

Tal como se ha explicado anteriormente, una optimización con restricciones se define con un parámetro $\alpha > (0,1)$, ya que no se permite que ese parámetro esté muy cerca de 0. Los criterios de optimización pueden ser definidos como:

45 A. Error mínimo ponderado de interpolación lineal:

$$\min \left(\sum_{i=0}^{1763} W_i (A_i - A'_i)^2 \right)$$

con W_i factor de ponderación, definido, por ejemplo, en función de $|A_i - A'_i|$ o, por ejemplo, de $|A_i|$. Los factores de ponderación pueden ser usados para aumentar el efecto de los errores en relación al error de aproximación o al valor absoluto de las muestras.

B. Espectro de frecuencia similar.

Se define $F_A = \text{DFT}(A)$ y $F_{A'} = \text{DFT}(A')$ como la transformada discreta de Fourier de las secciones de audio A con A_i ($0 \leq i < 1.764$) y A' con A'_i ($0 \leq i < 1.764$). F_{A_i} y $F_{A'_i}$ son números complejos. El espectro de potencia se define por $PF_{A_i} = F_{A_i} * \text{cj}(F_{A_i})$ y $PF_{A'_i} = F_{A'_i} * \text{cj}(F_{A'_i})$, siendo $\text{cj}()$ el complejo conjugado. PF_{A_i} ($0 \leq i < 882$) es el espectro de potencia de la sección A para una frecuencia entre 0 y 22,05 kHz. $PF_{A'_i}$ ($0 \leq i < 882$) es el espectro de potencia de la sección A'.

Cuando se optimizan los parámetros del filtro (α , β , γ , δ), los criterios pueden ser definidos en base al espectro de potencia del flujo aproximado y el flujo original. Debido a que el nuevo flujo se crea copiando cada muestra par en el nuevo flujo, se puede esperar que el espectro de potencia para frecuencias en el intervalo de 0 - 11,025 kHz será similar. Sin embargo, el espectro de potencia para frecuencias en el intervalo de 11,025 kHz hasta 22,050 kHz puede ser sustancialmente diferente. Por tanto, los criterios de optimización podrían definirse como:

$$\min \left(\sum_{i=434}^{868} W_i (PF_{A_i} - PF_{A'_i})^2 \right)$$

con W_i factores de ponderación, por ejemplo, para aumentar el efecto de los errores de potencia a frecuencia elevada. En el ejemplo, obsérvense los errores de potencia para el intervalo de frecuencias de 10,85 kHz (434) a 21,70 kHz (868).

C. Pueden aplicarse criterios de optimización más complejos, como aquellos que tienen en cuenta las características perceptuales de audio, o una combinación de diferentes criterios. En cualquier caso, se obtiene un conjunto de parámetros (α , β , γ , δ) a ser usados por el decodificador y el codificador. Los esquemas del codificador y del decodificador son independientes de los criterios de optimización seleccionados para los parámetros del filtro.

g. Implementación con enteros.

El procedimiento de codificación según las realizaciones de la presente invención requiere varias etapas. Se parte de los flujos A y B originales y, a continuación, los parámetros del filtro son optimizados para las secciones de 40 ms. Usando estos filtros, A' y B' se obtienen como flujos aproximados. A continuación, estos flujos A' y B' aproximados son mezclados según un esquema de mezclado seleccionado. Se definen los factores de compresión (representados por los flujos de compresión) y éstos se aplican a los flujos A & B originales, antes de la mezcla, para obtener A'' y B''.

Tal como se ha explicado en los párrafos anteriores, el decodificador y el codificador serán implementados usando operaciones con números enteros, para hacer que la codificación y la decodificación sean independientes de la implementación aritmética en coma flotante. Debido a este enfoque, se introducen errores de redondeo. Esto es especialmente cierto para fórmulas como: $A^{*}_{2i+5} = \alpha.A^{*}_{2i+6} + \beta.A^{*}_{2i+4} + \gamma.A^{*}_{2i+2} + \delta.A^{*}_{2i}$ o $A^{*}_{2i+6} = (1/\alpha).(A^{*}_{2i+5} - \beta.A^{*}_{2i+4} - \gamma.A^{*}_{2i+2} - \delta.A^{*}_{2i})$.

Tal como se ha indicado anteriormente, los parámetros (α , β , γ , δ) del filtro se representan como números enteros (A, B, C, D) de 16 bits, con números de base (A_b , B_b , C_b , D_b). Usando esta representación, la fórmula de codificación:

$$A^{*}_{2i+5} = \alpha.A^{*}_{2i+6} + \beta.A^{*}_{2i+4} + \gamma.A^{*}_{2i+2} + \delta.A^{*}_{2i}$$

se convierte en una operación con enteros de 32 bits:

$$A^{*}_{tmp} = (1 \ll A_b).A.A^{*}_{2i+6} + (1 \ll B_b).B.A^{*}_{2i+4} + (1 \ll C_b).C.A^{*}_{2i+2} + (1 \ll D_b).D.A^{*}_{2i}$$

y es convertida de nuevo a un valor entero de 16 bits:

Si $(A^{*}_{tmp} > 0)$ entonces $A^{*} = (A^{*}_{tmp} + 16.384) / 32.768$;

Si no $A^{*} = (A^{*}_{tmp} - 16.384) / 32.768$;

Esta implementación con enteros de esta aproximación, como tal, puede ser otra fuente generadora de condiciones de rebasamiento, que necesita ser comprobada de nuevo:

Si $(|A^{*}| > \text{UmbralMax})$ entonces "Informa acerca de condición de rebasamiento!!"

Si no $A^{*}_{2i+5} = A^{*}_{tmp}$

En el caso de que se produzca una nueva condición de rebasamiento, los flujos de compresión iniciales deben ser redefinidos, usando valores UmbralMax más pequeños.

Después de analizar los efectos de la implementación con enteros de la fórmula de codificación, los efectos de la implementación del decodificador con enteros deben ser analizados también. La fórmula de decodificación usada es:

$$A''_{2i+6} = (1/\alpha) \cdot (A''_{2i+5} - \beta \cdot A''_{2i+4} - \gamma \cdot A''_{2i+2} - \delta \cdot A''_{2i}).$$

Convertida a una fórmula con números enteros de 32 bits:

$$5 \quad A''_{tmp} = (32.768 \cdot A''_{2i+5}) - (1 \ll Bb) \cdot B \cdot A''_{2i+4} - (1 \ll Cb) \cdot C \cdot A''_{2i+2} - (1 \ll Db) \cdot D \cdot A''_{2i}$$

$$A'' = A''_{tmp} / ((1 \ll Ab) \cdot A)$$

Una vez más, esta implementación con números enteros de esta aproximación puede ser otra fuente generadora de condiciones de rebasamiento, que debe ser revisada de nuevo:

Si $(|A''| > \text{UmbralMax})$ entonces 'Informa acerca de condición de rebasamiento!!'

10 Además, debido a las conversiones a enteros y al error de redondeo asociado, A'' no es necesariamente igual a A''_{2i+6} . Para garantizar que el decodificador funcionará correctamente, el valor A''_{2i+6} original es "actualizado" o "reemplazado" con el valor aproximado de A'' . Este valor actualizado de A''_{2i+6} será usado, a continuación, como uno de los valores usados en la siguiente aproximación de A''_{2i+7} .

15 Una consecuencia de estos errores de redondeo es que no sólo se aproximan las muestras impares del flujo, sino también los valores pares son "actualizados" con los valores para los cuales la fórmula de decodificación proporcionará el "valor" (actualizado) correcto. Pueden necesitarse actualizaciones sucesivas para las muestras pares, lo que puede conducir a una divergencia. Por lo tanto, se requiere indicar el nivel o las correcciones de los errores de redondeo. Tal como se ha indicado anteriormente, se realiza una optimización con restricciones para los parámetros $(\alpha, \beta, \gamma, \delta)$ para evitar grandes errores de redondeo. Para el parámetro α , se establece, típicamente, un límite inferior de 0,1, que
20 representa un buen equilibrio entre las características de optimización del filtro, por un lado, y la minimización de los errores de redondeo, por otro. Sin embargo, si el efecto de los errores de redondeo es aún demasiado alto, la optimización debe repetirse con mayores restricciones para los parámetros.

En un aspecto adicional, los datos MIDI y las letras de la presente invención están incluidos en el formato del codificador.

25 Tal como se ha explicado en las secciones anteriores, el formato CD+G(w) usado para las realizaciones de la presente invención incorpora también datos MIDI, así como letras. Las letras se incluyen como parte del flujo de datos MIDI.

MIDI (Musical Instruments Digital Interface) define un protocolo serie asíncrono, usado por los instrumentos musicales para transmitir datos. La tasa de bits se define como de 31.250 bits/seg. Un byte es transferido usando un bit de inicio y parada, y la máxima tasa de bytes real es: 3.125 bytes por segundo.

30 CD+G(w) utiliza, junto con los datos de audio PCM, canales de datos adicionales (canales de sub-código R-W desde el CD) para almacenar parámetros adicionales. Un sector de datos de audio PCM (esto es, 1/75-ésima parte de un segundo) tiene 72 bytes de sub-código R-W adicionales. Debido a una capa de corrección de errores, y algunas estructuras lógicas de datos de cabecera, el modo CD+G USER permite almacenar sólo 48 bytes adicionales, o $48 \times 75 = 3.600$ bytes por segundo. Se usan 1.800 bytes por segundo para los parámetros del filtro de audio, tal como se ha explicado en las secciones anteriores, que toman 24 bytes por sector de los sub-códigos CD. Esto deja sólo 1.800 bytes libres para estos
35 datos MIDI u otros 24 bytes por sección. Por lo tanto, se concluye que este ancho de banda ($75 \times 24 = 1.800$ bytes por segundo) no permite almacenar datos MIDI en un "formato de emisión en tiempo real".

40 Por lo tanto, los datos MIDI usados en el formato CD+G(w) comprenden datos discretos con la información de marca de tiempo incluida. Los datos discretos son similares al evento o eventos de datos MIDI y la marca de tiempo es el tiempo delta, relativo al inicio de esa sección de datos de sub-código particular. Un mensaje MIDI típico comprende 2 ó 3 bytes. Por ejemplo:

Mensaje Nota ON: Byte 1: Nota On MIDI (0x90) en el canal 2 (| 0x02)

Byte 2: Número de nota MIDI 64 (0x40)

Byte 3: Velocidad nota On MIDI 32 (0x20)

Mensaje Nota OFF: Byte 1: Nota Off MIDI (0x80) en el canal 2 (| 0x02)

45 Byte 2: Número nota MIDI 64 (0x40)

Byte 3: Velocidad nota Off MIDI 0 (0x0)

Mensaje Nota Off (estado de ejecución): Byte 1: Número de nota MIDI 64 (0x40)

Byte 2: Velocidad nota Off MIDI 0 (0x0)

5 Un mensaje MIDI típico requiere 3 bytes, o 30 bits, de manera que la precisión de temporización no es mejor que 1 ms. En CD+G(w), se usa 1 byte para definir la temporización delta relativa al inicio de la sección sub-código de CD que contiene los mensajes MIDI. Una sección es 1/75-ésima parte de un segundo o 13,3 ms. Este periodo de tiempo es dividido en 49 marcas de tiempo, con una precisión del intervalo de marca de tiempo (12/44.100) o aprox. 272 μ s. Como tal, la resolución temporal del CD+G(w) es 4 veces mejor en comparación con MIDI. Debido a que se usan valores de 8 bits como marca de tiempo, pueden definirse retrasos para los eventos MIDI de aproximadamente 70 ms relativos al inicio de la sección que contiene el mensaje MIDI. Esto permite agrupar, de manera más efectiva, los mensajes MIDI dentro de un intervalo de tiempo más corto.

10 MIDI restringe la definición de velocidad a valores de 7 bits, pero CD+G(w) añade 1 bit extra para aumentar la resolución de la velocidad. Los mensajes MIDI típicos son de 3 bytes o de 2 bytes. Por cada sección de Sub-código CD hay 24 bytes disponibles. Un mensaje MIDI de 3 bytes requiere 4 bytes de datos de sub-código de CD (marca de tiempo incluida), mientras que un mensaje MIDI de 2 bytes requiere 3 bytes. O, en 24 bytes pueden definirse 6x 'mensaje MIDI de 3 bytes' [6x (3+1)] u 8x 'mensaje MIDI de 2 bytes' [8x (2+1)]. Si se toma la media, aproximadamente 17 bytes de datos MIDI efectivos están disponibles dentro de 1 sector del sub-código CD+G(w), o $17 \times 75 = 1.275$ bytes por segundo.

Una comparación de los datos MIDI estándar, y los datos MIDI incorporados en el formato CD+G(w), según una realización de la presente invención, se puede encontrar en la Figura 9.

20 Se usará MIDI en el formato CD+G(w), sólo como datos para información típica de instrumento solista y/o letras, y no para remplazar una actuación de una banda o una orquesta completa. Debido a este uso restringido, el ancho de banda limitado de 1.275 bytes por segundo, no se considera un problema. Además, se gana en precisión de temporización y de velocidad, que es lo que se necesita, específicamente, para las aplicaciones de reproducción de instrumentos (reproducción automatizada en grupo).

25 Finalmente, una suma de control CRC (Código de redundancia cíclica) de 16 bits, es decir, 16 bits añadidos al final de un bloque de datos para fines de comprobación de errores, con una cabecera de 8 bits, podría ser añadida por cada 3 sectores (40 ms) como parte de los datos MIDI de los datos de subcódigo RW, abarcando tanto el PCM como el subcódigo RW, como una referencia usada por el decodificador para verificar la integridad tanto de PCM como de los datos de subcódigo. Si se detecta un error de CRC, el decodificador debería decidir no decodificar el audio PCM, reproducir como estéreo regular y gestionar los eventos de datos MIDI en una manera apropiada.

30 De lo expuesto anteriormente, se apreciará que la presente invención se refiere a un codificador para mezclar una pluralidad de canales independientes de audio mono en una grabación estéreo y generar un conjunto restringido de parámetros adicionales usados para la masterización de una pista de audio de un dispositivo de almacenamiento, en el que cada canal de audio comprende un primer número de muestras por unidad de tiempo, estando almacenada la pluralidad de canales independientes de audio mono en el dispositivo de almacenamiento con menos que el número de muestras por unidad de tiempo de cada canal de audio con la adición de los valores calculados, calculados usando un filtro de interpolación con los parámetros del filtro, para las muestras intermedias, en las que las muestras intermedias calculadas se aproximan a las muestras intermedias de los canales de audio, estando almacenados los parámetros del filtro en el dispositivo de almacenamiento en los parámetros adicionales.

40 La presente invención se refiere también a un decodificador para decodificar una pluralidad de canales de audio mono grabados en una grabación en estéreo usando un conjunto limitado de parámetros adicionales para la masterización de una pista de audio de un dispositivo de almacenamiento, en el que cada canal de audio se reconstruye a partir de un primer número de muestras por unidad de tiempo, la pluralidad de canales de audio mono están almacenados en el dispositivo de almacenamiento con menos que el número de muestras por unidad de tiempo de cada canal de audio con la adición de los valores calculados, calculados usando un filtro de interpolación con los parámetros del filtro, para muestras intermedias, las muestras intermedias calculadas se aproximan a las muestras intermedias de los canales de audio, siendo almacenados los parámetros del filtro en el dispositivo de almacenamiento en los parámetros adicionales.

45 La presente invención se refiere también a un dispositivo de almacenamiento que ha almacenado una pluralidad de canales de audio mono grabados en una grabación en estéreo usando un conjunto restringido de parámetros adicionales para la masterización de una pista de audio del dispositivo de almacenamiento, en el que cada canal de audio puede ser reconstruido a partir de un primer número de muestras por unidad de tiempo, estando almacenada la pluralidad de canales de audio mono en el dispositivo de almacenamiento con menos que el número de muestras por unidad de tiempo de cada canal de audio con la adición de los valores calculados, calculados usando un filtro de interpolación con los parámetros del filtro, para las muestras intermedias, las muestras intermedias calculadas se aproximan a las muestras intermedias de los canales de audio, siendo almacenados los parámetros del filtro en el dispositivo de almacenamiento en

los parámetros adicionales.

- En un aspecto, el codificador/decodificador según la presente invención puede ser usado para codificar y mezclar un número de canales independientes, por ejemplo, 4 x canales de 44,1 KHz, de 16bit, mono, en una grabación de audio en estéreo, de 44,1 KHz, de 16 bits, que genera un conjunto de parámetros de datos adicionales. La grabación puede ser realizada sobre cualquier medio de grabación adecuado, tal como un dispositivo de memoria de estado sólido o un disco óptico. Por ejemplo, estos parámetros de datos adicionales son copiados a un disco óptico de audio estándar, tal como un CD-DA (Audio Digital) usando, por ejemplo, los canales de sub-código del formato CD+G USER. El decodificador es capaz de regenerar la pluralidad de canales independientes. Se proporcionan diferentes características para su uso en diversas aplicaciones de audio, tales como Sing-Along & Play-Along o Automated Play-Along.
- 5 Un conjunto de parámetros de datos es generado para cada período de tiempo, por ejemplo, para cada 40 ms de grabación de audio, y pueden incluir
- Parámetros del filtro decodificador, para la reconstrucción de los canales de audio originales
 - Parámetros y modo de mezclado de audio, para indicar la presencia de datos MIDI y letras
 - Parámetros de atenuación y panoramización de audio
- 10 - Datos de panoramización dinámica de audio 3D
- Parámetros del filtro de audio para una extensión de tiempo y/o cambio de tono mejorados
 - Datos MIDI para instrumentos musicales electrónicos o automatizados
 - Datos de letras (integrados en MIDI)
 - Datos de suma de control CRC de 16 bits, con cabecera de 8 bits, (integrado en MIDI)
- 15 La presente invención puede ser usada con grabaciones de audio y/o vídeo en cualquier forma de dispositivo de almacenamiento, por ejemplo, de estado sólido, discos ópticos, tales como CD-DA, CD-ROM, DVD-AUDIO o DVD-ROM, cinta, tales como cinta DAT. También, más de 4 canales de audio pueden ser grabados de manera que cada uno puede ser des-mezclado, según sea necesario, por ejemplo, usando menos muestras de cada pista de audio mono y calculando más puntos intermedios. Sin embargo, esto puede reducir la calidad global del audio y, normalmente, 4 pistas son
- 20 suficientes.
- 25

REIVINDICACIONES

1. Codificador para mezclar una pluralidad de canales de audio mono independientes en una grabación de audio que comprende muestras mezcladas para su almacenamiento en una pista de audio de un dispositivo de almacenamiento, en el que cada canal de audio mono comprende un primer número de muestras por unidad de tiempo, cada canal de audio mono tiene el mismo número de muestras por unidad de tiempo, y el mismo tiempo de referencia, la señal de audio comprende menos muestras mezcladas por unidad de tiempo que una suma del número de muestras por unidad de tiempo de todos los canales de audio mono usados en la señal de audio,
5 en el que el codificador está adaptado para:
10 derivar muestras aproximadas filtrando muestras de un flujo de audio de los canales de audio mono respectivos usando un filtro de interpolación, en el que dichas muestras aproximadas tienen temporizaciones intermedias con respecto a dichas muestras del flujo de audio,
15 proporcionar un flujo de muestras mezcladas, sumando una primera muestra de flujo de audio de un primer canal de audio mono y una primera muestra aproximada derivada a partir de un segundo canal de audio mono para obtener una primera muestra mezclada, en el que las primeras muestras corresponden a una primera temporización y sumando una segunda muestra de flujo de audio del segundo canal de audio mono y una segunda muestra aproximada derivada a partir del primer canal de audio mono para obtener una segunda muestra mezclada, en el que las segundas muestras corresponden a una segunda temporización sucesiva a la primera temporización
20 en el que el codificador está adaptado además para:
25 generar un conjunto restringido de parámetros adicionales para la masterización de la pista de audio del dispositivo de almacenamiento que comprende la grabación de la señal de audio, en el que el conjunto restringido de parámetros comprende parámetros del filtro que comprenden muestras que permiten calcular la primera muestra del primer canal de audio mono y la segunda muestra del segundo canal de audio mono de la primera muestra mezclada.
2. Codificador según la reivindicación 1, en el que el dispositivo de almacenamiento puede ser reproducido usando un reproductor de disco óptico de audio, de manera que, en un primer modo, toda la pluralidad de canales de audio mono son reproducidos como la grabación de señal de audio mezclada y, en un segundo modo, a partir de la grabación de señal de audio, al menos uno de entre la pluralidad de canales de audio mono puede ser des-
30 mezclado y la grabación de señal de audio es reproducida como una grabación de audio cuadrafónica verdadera.
3. Codificador según la reivindicación 1 ó 2, en el que el codificador está adaptado para derivar las muestras aproximadas por medio de interpolación lineal.
4. Codificador según la reivindicación 1 ó 2, en el que el codificador está adaptado para derivar las muestras aproximadas usando un filtro FIR.
- 35 5. Codificador según la reivindicación 1, 2, 3 ó 4, en el que el dispositivo de almacenamiento es un disco óptico.
6. Codificador según la reivindicación 5, en el que el disco óptico es compatible con Redbook.
7. Codificador según la reivindicación 5, en el que el conjunto de parámetros adicionales son almacenados dentro del formato CD+G.
- 40 8. Codificador según cualquiera de las reivindicaciones 6 ó 7, en el que el conjunto restringido de parámetros adicionales no excede una capacidad de almacenamiento disponible como datos sub-código dentro del formato CD+G.
9. Codificador según cualquiera de las reivindicaciones anteriores, en el que los canales de audio son canales de audio PCM.
- 45 10. Codificador según cualquiera de las reivindicaciones anteriores, que está adaptado para proporcionar parámetros de filtro adicionales en la grabación de señal de audio para des-mezclar el audio del dispositivo de almacenamiento que comprende la grabación de señal de audio.
11. Codificador según cualquiera de las reivindicaciones anteriores, que está adaptado para generar los parámetros del filtro para des-mezclar la pista de audio del dispositivo de almacenamiento que comprende la

grabación de la señal de audio, realizando una optimización de los parámetros del filtro.

12. Decodificador para decodificar un primer canal de audio mono y un segundo canal de audio mono a partir de una grabación de audio que comprende un flujo de muestras mezcladas almacenadas en una pista de audio de un dispositivo de almacenamiento desde un codificador según la reivindicación 1, en el que la grabación de audio comprende además un conjunto restringido de parámetros adicionales, en el que cada canal de audio mono comprende un primer número de muestras por unidad de tiempo, cada canal de audio mono tiene el mismo número de muestras por unidad de tiempo, y el mismo tiempo de referencia y la grabación de señal de audio comprende menos muestras mezcladas por unidad de tiempo que una suma del número de muestras por unidad de tiempo del primer canal de audio mono y del segundo canal de audio mono,
- estando adaptado el decodificador para procesar el flujo de muestras mezcladas:
 - restando una primera muestra decodificada previamente de un primer canal de audio mono de una primera muestra mezclada, en el que la primera muestra decodificada previamente es derivada a partir de una muestra conocida de los parámetros del filtro almacenados en el conjunto restringido de parámetros adicionales, correspondiendo las primeras muestras a una primera temporización para obtener una primera muestra del segundo canal de audio mono, y restando una segunda muestra decodificada previamente de una segunda muestra mezclada para obtener una segunda muestra aproximada del primer canal de audio mono, en el que las segundas muestras corresponden a una segunda temporización,
 - calculando una tercera muestra a partir de un primer canal de audio mono a partir de una primera muestra decodificada previamente y una segunda muestra aproximada del primer canal de audio mono, y
 - decodificando una tercera muestra aproximada subsiguiente del segundo canal de audio mono usando la tercera muestra calculada de un primer canal de audio mono y una tercera muestra mezclada, en el que las terceras muestras corresponden a una tercera temporización, en el que las temporizaciones primera, segunda y tercera tienen temporizaciones sucesivas.
13. Decodificador según la reivindicación 12, en el que la grabación de señal de audio puede ser reproducida, en un primer modo, cuando se lee solo la grabación de señal de audio y, en un segundo modo, leyendo, de manera complementaria, el conjunto restringido de parámetros adicionales desde el dispositivo de almacenamiento, al menos uno de entre la pluralidad de canales de audio mono puede ser des-mezclado y la grabación de señal de audio puede ser reproducida con al menos un canal de audio eliminado y no reproducido, reproducida como una grabación de audio cuadrafónica verdadera.
14. Decodificador según la reivindicación 12 ó 13, en el que el decodificador está adaptado para calcular las muestras aproximadas mediante interpolación lineal.
15. Decodificador según la reivindicación 12 ó 13, en el que el decodificador está adaptado para calcular las muestras aproximadas usando un filtro FIR.
16. Decodificador según la reivindicación 12, 13, 14 ó 15, en el que el dispositivo de almacenamiento es un disco óptico.
17. Decodificador según la reivindicación 16, en el que el disco óptico es compatible con Redbook.
18. Decodificador según cualquiera de las reivindicaciones 12 ó 13, en el que el conjunto de parámetros adicionales son recuperados desde el formato CD+G.
19. Decodificador según cualquiera de las reivindicaciones anteriores, en el que los canales de audio son canales de audio PCM.
20. Dispositivo de almacenamiento que comprende una grabación de señal de audio obtenida desde un codificador según se reivindica en cualquiera de las reivindicaciones 1 a 11.
21. Sistema de audio que comprende un codificador según cualquiera de las reivindicaciones 1 a 11, un decodificador según cualquiera de las reivindicaciones 12 a 19 y un dispositivo de almacenamiento según la reivindicación 20.
22. Sistema de audio según la reivindicación 21, en el que el sistema soporta la reproducción de audio cuadrafónico verdadero.
23. Sistema de audio según la reivindicación 21 ó 22, en el que cuando es usado en aplicaciones de sonido

envolvente, el decodificador es capaz de reconstruir los canales de audio posteriores en estéreo completamente aislados de los canales frontales en estéreo, a partir de un dispositivo de almacenamiento codificado cuadrafónicamente.

- 5 24. Sistema de audio según la reivindicación 21 ó 22, en el que cuando es usado en aplicaciones de sonido envolvente, el decodificador es capaz de panoramizar dinámicamente un canal de audio mono a partir de un disco óptico de un dispositivo de almacenamiento codificado dentro de un espacio auditivo 3D, para reproducir una fuente de sonido en movimiento similar a cuando un artista se mueve en el escenario.
- 10 25. Sistema de audio según la reivindicación 21 ó 22, en el que el decodificador está adaptado para ejecutar la operación inversa del codificador y tanto el codificador como el decodificador están implementados en aritmética de números enteros, siendo independientes de una implementación en coma flotante.
26. Sistema de audio según cualquiera de las reivindicaciones 21 a 25, en el que el codificador está dispuesto para realizar una aproximación de los canales de audio mono previamente a mezclar estos canales, para reducir errores introducidos por una conversión subsiguiente a números enteros.
- 15 27. Sistema de audio según la reivindicación 25, que está adaptado además para la conversión a aritmética de números enteros, estando adaptados el codificador y/o decodificador para compensar los errores de redondeo y conversión para proporcionar un procedimiento de decodificación que es la operación inversa de la codificación.
28. Sistema de audio según la reivindicación 21 ó 22, en el que los parámetros del filtro de audio requeridos por el decodificador y que son una parte del conjunto restringido de parámetros adicionales, son regenerados para cada periodo de tiempo, que es corto en comparación con la resolución de audio del oído humano.
- 20 29. Sistema de audio según la reivindicación 28, en el que el periodo de tiempo es de 40 ms de audio.
30. Sistema de audio según la reivindicación 21 ó 22, en el que el codificador incorpora un compresor dinámico que previene rebasamientos cuando se mezclan y/o codifican la pluralidad de canales mono en un canal de audio.

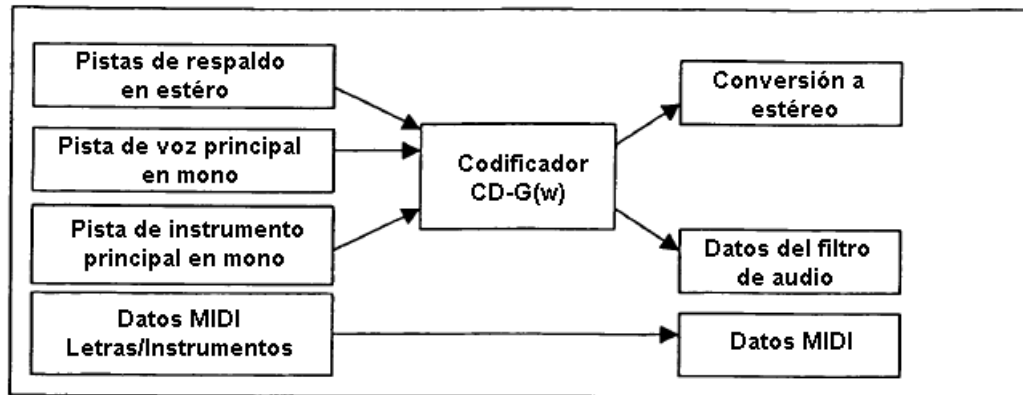


Figura 1

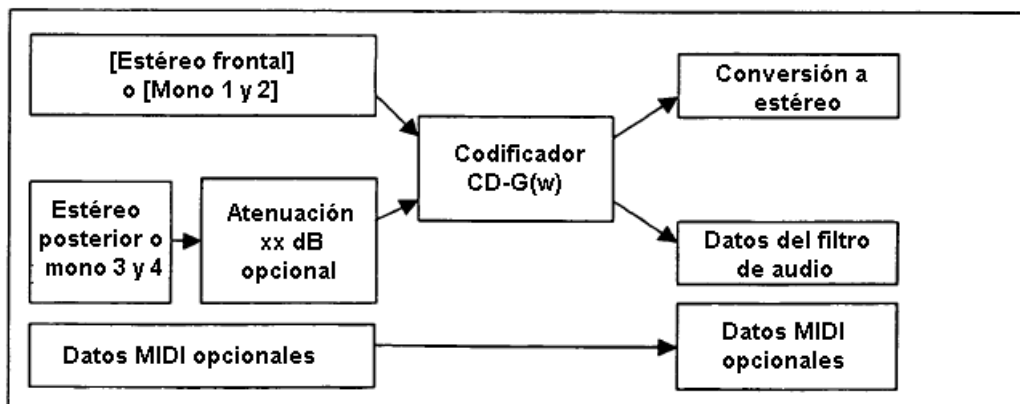


Figura 2

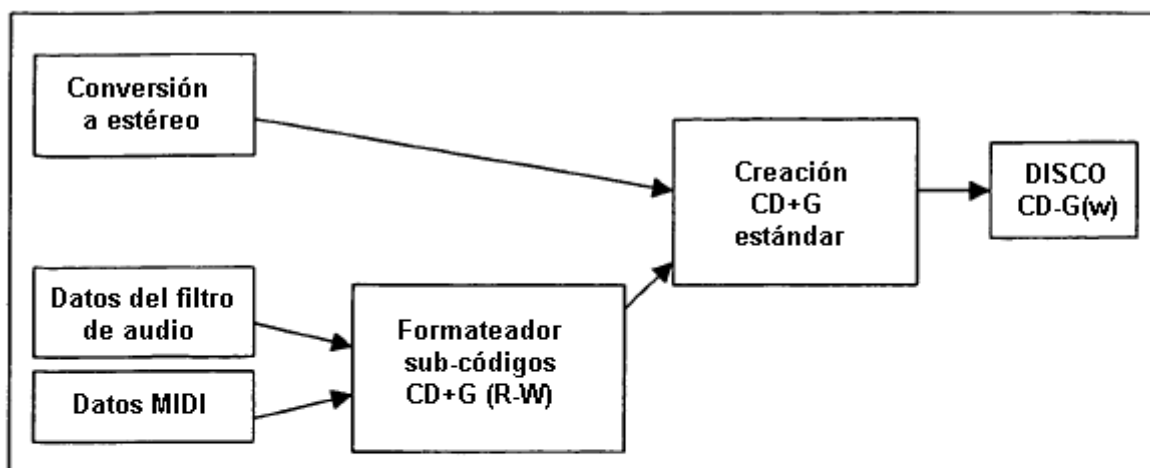


Figura 3

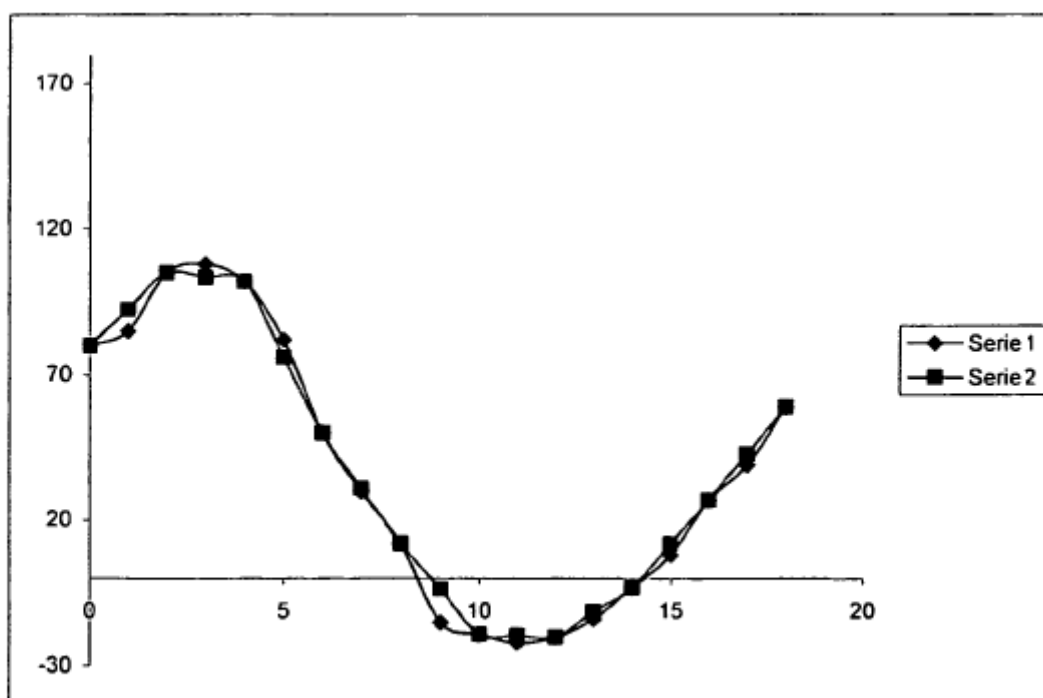


Figura 4

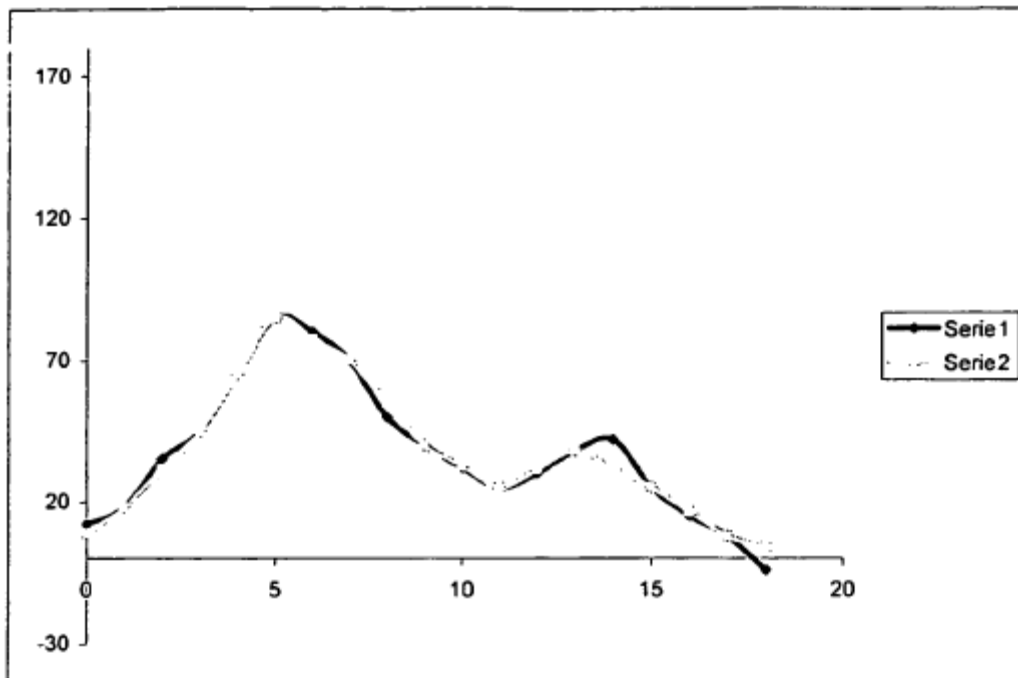


Figura 5

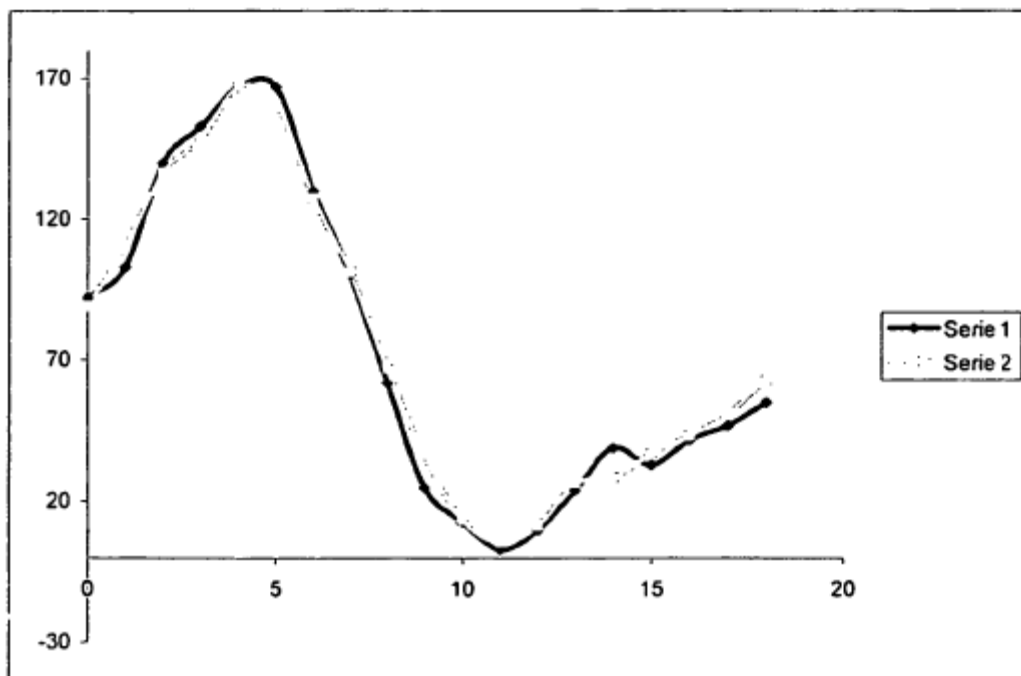


Figura 6

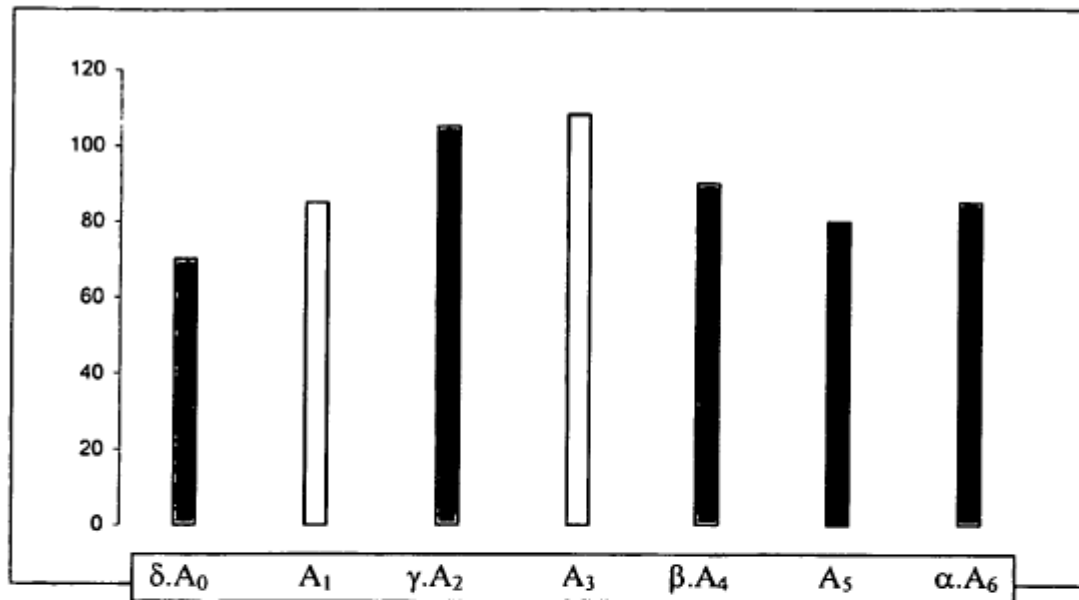


Figura 7

Parámetro	Datos MIDI estándar	Datos MIDI CD+G(w)	Compar.
Tasa (Byte) efectiva	3.125	1.275	60% menos
Precisión temporización	1 ms	272 us	4x mejor
Resolución de velocidad	128 niveles	256 niveles	2x mejor

Figura 9

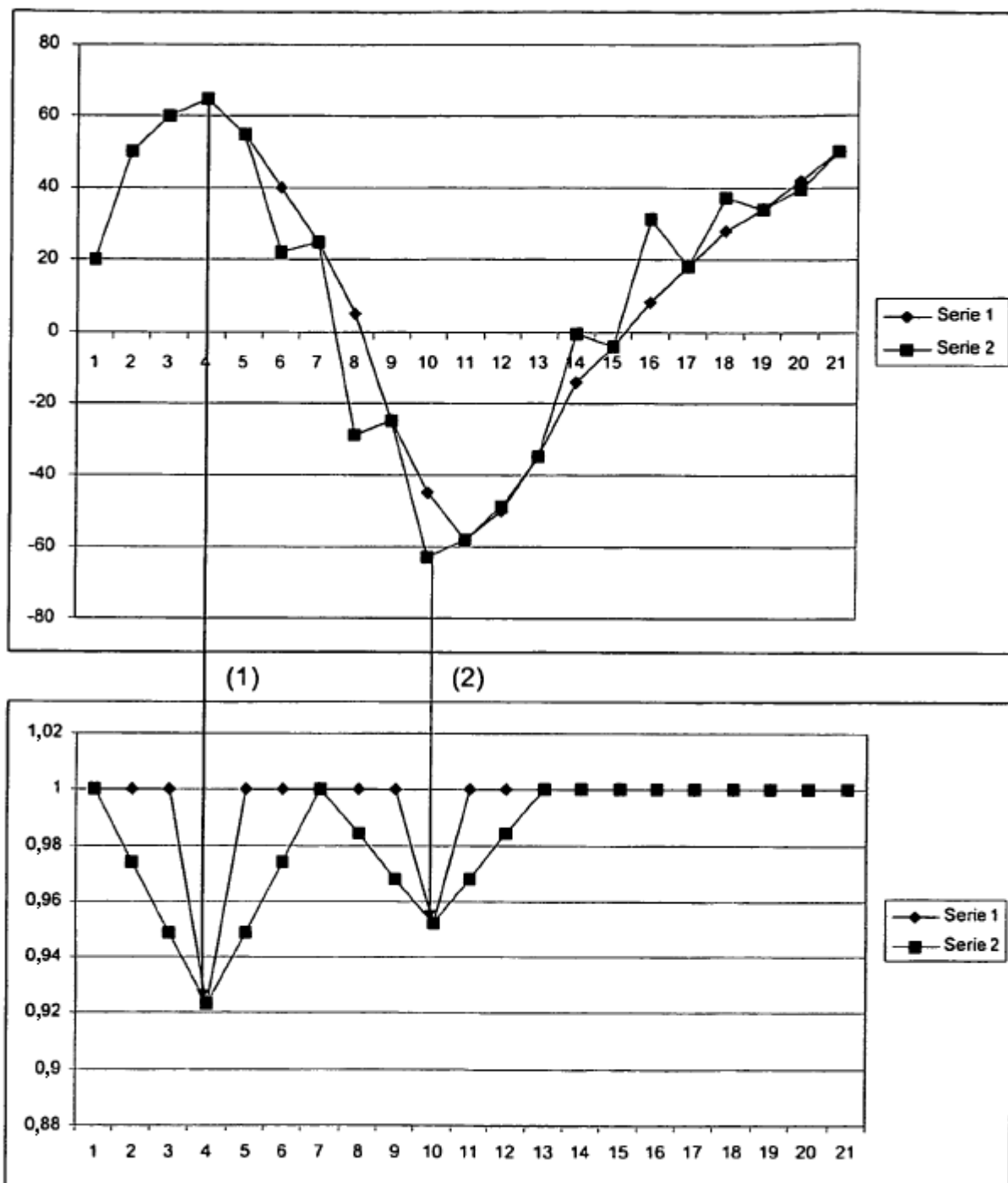


Figura 8