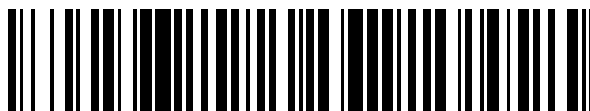


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 731 428**

51 Int. Cl.:

**G10L 19/008** (2013.01)

**G10L 19/20** (2013.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **04.03.2015 PCT/US2015/018804**

87 Fecha y número de publicación internacional: **24.09.2015 WO15142524**

96 Fecha de presentación y número de la solicitud europea: **04.03.2015 E 15764758 (7)**

97 Fecha y número de publicación de la concesión europea: **08.05.2019 EP 3120346**

54 Título: **Codificación residual en un sistema de audio basado en objetos**

30 Prioridad:

**20.03.2014 US 201461968111 P**  
**12.02.2015 US 201514620544**

45 Fecha de publicación y mención en BOPI de la  
traducción de la patente:  
**15.11.2019**

73 Titular/es:

**DTS, INC. (100.0%)**  
**5220 Las Virgenes Road**  
**Calabasas, CA 91302, US**

72 Inventor/es:

**KALKER, ANTONIUS y**  
**SEROUSSI, GADIEL**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

ES 2 731 428 T3

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).

## DESCRIPCIÓN

Codificación residual en un sistema de audio basado en objetos

## 5 Solicitudes relacionadas:

Esta solicitud reivindica prioridad a la Solicitud Provisional de Estados Unidos con N.º de Serie 61/968.111, presentada el 20 de marzo de 2014, y titulada "Residual Encoding in an Object-Based Audio System", y a la Solicitud No Provisional de Estados Unidos N.º 14/620.544, presentada el 12 de febrero de 2015, y titulada "Residual Encoding In An Object-Based Audio System".

## Antecedentes de la invención:

### 1. Campo de la invención:

Esta invención se refiere en general a la compresión y descompresión de audio de múltiples canales con pérdidas en general, y más específicamente a compresión y descompresión de señales de audio de múltiples canales mezcladas de manera descendente de una manera que facilita la mezcla ascendente de las señales de audio de múltiples canales recibidas y descomprimidas.

### 2. Descripción de la técnica relacionada:

Los sistemas de entretenimiento de audio y audio-visuales han progresado desde comienzos humildes, pudiendo reproducir audio monoaural a través de un único altavoz. Los sistemas de sonido envolvente modernos pueden grabar, transmitir y reproducir una pluralidad de canales, a través de una pluralidad de altavoces en un entorno de oyente (que puede ser un cine público o un "cine en casa" más privado). Hay disponible diversas disposiciones de altavoces de sonido envolvente: van por designaciones tales como "envolvente 5.1", "envolvente 7.1", e incluso envolvente 20.2 (donde el número a la derecha del punto decimal indica un canal de efectos de baja frecuencia). Para cada configuración de este tipo, son posibles diversas disposiciones físicas de altavoces; pero en general los mejores resultados se conseguirán si la geometría de representación es similar a la geometría supuesta por los ingenieros de audio que mezclaron y elaboraron el master de los canales grabados.

Puesto que son posibles diversos entornos de representación y geometrías más allá de la predicción de los ingenieros de mezcla, y puesto que el mismo contenido puede reproducirse en diversas configuraciones de escucha o entornos, la multiplicidad de configuraciones de sonido envolvente presenta numerosos desafíos al ingeniero o artista que desea entregar una experiencia de escucha fidedigna. Puede emplearse cualquiera de un enfoque "basado en canal" o (más recientemente) un enfoque "basado en objeto" para entregar la experiencia de escucha de sonido envolvente.

En un enfoque basado en canal, cada canal se graba con la intención de que el sonido se representará durante la reproducción en un altavoz correspondiente. La disposición física de los altavoces pretendidos se predetermina o al menos se supone de manera aproximada durante la mezcla. En contraste, en un enfoque basado en objeto se graba, almacena y transmite una pluralidad de objetos de audio independientes de manera separada, conservando su relación síncrona, pero independientemente de cualesquiera presuposiciones acerca de la configuración o geometría de los altavoces o entorno de reproducción pretendidos. Ejemplos de objetos de audio serían un único instrumento musical, una sección de conjunto tal como una sección de viola considerada como una voz musical unitaria, una voz humana o un efecto de sonido. Para conservar las relaciones espaciales, los datos digitales que representan los objetos de audio incluyen para cada objeto ciertos datos ("metadatos") que simbolizan información asociada con la fuente de sonido particular: por ejemplo, la dirección de vector, proximidad, volumen, movimiento y extensión de la fuente de sonido pueden codificarse simbólicamente (preferentemente de una manera apta para variación de tiempo) y esta información se transmite o graba junto con la señal de sonido particular. La combinación de una forma de onda de fuente de sonido independiente y los metadatos asociados juntos comprende un objeto de audio (almacenado como un fichero de objeto de audio). Este enfoque tiene la desventaja de que puede presentarse de manera flexible, en muchas configuraciones diferentes; sin embargo, se pone la carga en el procesador de representación ("motor") para calcular la mezcla apropiada basándose en la geometría y configuración de los altavoces y entorno de reproducción.

En ambos enfoques basado en canal y basado en objeto para audio, de manera frecuente es deseable transmitir una señal mezclada de manera descendente (A más B) de tal manera que los dos canales independientes (y objetos, A y B) pueden separarse ("mezclados de manera ascendente") durante la reproducción. Una motivación para transmitir una mezcla descendente debe ser mantener la compatibilidad hacia atrás, para que un programa mezclado de manera descendente pueda reproducirse en monoaural, estéreo de dos canales convencional, o (más en general) en un sistema con menos altavoces que el número de canales u objetos en el programa grabado. Para recuperar la pluralidad más alta de canales u objetos, se emplea un proceso de mezcla ascendente. Por ejemplo, si se transmite la suma C de las señales A y B (A+B), y si también se transmite B, entonces el receptor puede construir fácilmente  $A = (A+B) - B$ . Como alternativa, se pueden transmitir señales compuestas (A+B) y (A-B), a continuación

recuperar A y B tomando combinaciones lineales de las señales compuestas transmitidas. Muchos sistemas de la técnica anterior usan variaciones de este enfoque de "mezcla de matriz". Estos son en cierto modo satisfactorios al recuperar canales u objetos discretos. Sin embargo, cuando se suman grandes números de canales o especialmente de objetos, se hace difícil reproducir adecuadamente objetos o canales discretos individuales sin artefactos o requisitos de ancho de banda alto poco prácticos. Puesto que el audio basado en objeto a menudo implica números muy altos de objetos de audio independientes, hay implicadas enormes dificultades en la mezcla de manera ascendente de manera eficaz para recuperar objetos discretos de señales mezcladas de manera descendente, particularmente donde está restringida la tasa de datos (o más en general, el ancho de banda). Un ejemplo del enfoque basado en objeto para audio puede hallarse en el documento WO 2014/023443 A1.

En la mayoría de los sistemas prácticos para transmisión o grabación de audio digital, será altamente deseable algún método de compresión de datos. La tasa de datos se somete siempre a alguna restricción, y se desea siempre para transmitir audio de manera más eficaz. Esta consideración se hace cada vez más importante cuando se emplea un gran número de canales como canales discretos o mezclados de manera ascendente. En la presente solicitud el término "compresión" hace referencia a métodos de reducción del requisito de datos para transmitir o grabar señales de audio, si el resultado es reducción de tasa de datos o tamaño de fichero. (Esta definición no debería confundirse con compresión de rango dinámico, que en ocasiones también se denomina como "compresión" en otros contextos de audio no relevantes en este punto).

Los enfoques anteriores para comprimir señales mezcladas de manera descendente generalmente adoptan uno de dos métodos: codificación sin pérdidas o descripción redundante. Cualquiera puede facilitar la mezcla ascendente después de la descompresión, pero ambos tienen desventajas.

Codificación sin pérdidas y con pérdidas:

Suponiendo que  $A, B_1, B_2, \dots, B_m$  son señales independientes (objetos), que se codifican en un *flujo de código* y se envían a un *representador*. El objeto distinguido  $A$  se denominará como el *objeto de base*, mientras que  $B = B_1, B_2, \dots, B_m$  se denominarán como *objetos regulares*. En un sistema de audio basado en objetos, estamos interesados en representar objetos simultánea pero independientemente, de modo que, por ejemplo, cada objeto se representaría en una localización espacial diferente. Es deseable compatibilidad hacia atrás: en otras palabras, requerimos que el flujo codificado sea interpretable por sistemas heredados que ni están basados en objeto ni reconocen objetos, o que son aptos para menos canales. Tales sistemas únicamente pueden representar el *objeto compuesto o canal*  $C = A + B_1 + B_2 + \dots + B_m$  de una versión  $E(C)$ , de  $C$  codificada (comprimida). Por lo tanto, requerimos que se transmita el flujo de código incluya  $E(C)$ , seguido por descripciones de los objetos individuales, que se ignoran por los sistemas heredados. Por lo tanto, el flujo de código puede consistir en  $E(C)$  seguido por las descripciones  $E(B_1), E(B_2), \dots, E(B_m)$  de los objetos regulares. El objeto de base  $A$  se recupera a continuación decodificando estas descripciones y estableciendo  $A = C - B_1 - B_2 - \dots - B_m$ . Debería observarse, sin embargo, que la mayoría de códecs de audio usados en la práctica son *con pérdidas*, que significa que la versión decodificada  $Q(X) = D(E(X))$  de un objeto codificado  $E(X)$  es únicamente una aproximación de  $X$ , y por lo tanto no necesariamente idéntica a ella. La precisión de la aproximación en general depende de la elección del códec y del ancho de banda (o espacio de almacenamiento) disponible para el flujo de código. Mientras que es posible una *codificación sin pérdidas*, es decir  $Q(X) = X$ , normalmente requiere ancho de banda o espacio de almacenamiento significativamente mayor que una codificación con pérdidas. Lo último, por otra parte, puede proporcionar aún una reproducción de alta calidad que puede ser perceptualmente indistinguible de la original.

Descripción redundante:

Un enfoque alternativo es incluir una codificación explícita de ciertos objetos privilegiados  $A$  en el flujo de código, que consistiría por lo tanto en  $E(C), E(A), E(B_1), E(B_2), \dots, E(B_m)$ . Suponiendo que  $E$  es con pérdidas, este enfoque es probable que sea más económico que usar una codificación con pérdidas, pero aún no es un uso eficaz del ancho de banda. El enfoque es redundante, puesto que  $E(C)$  está evidentemente correlacionada a los objetos codificados individualmente  $E(A), E(B_1), E(B_2), \dots, E(B_m)$ .

## Sumario de la invención

La invención proporciona un método de descompresión y mezcla ascendente de una señal de audio compuesta comprimida y mezclada de manera descendente con las características de la reivindicación 1 y un método de compresión de una señal de audio compuesta que comprende una señal de mezcla total con las características de la reivindicación 8. Las realizaciones de la invención se identifican en las reivindicaciones dependientes.

La compresión con pérdidas y transmisión de una señal compuesta mezclada de manera descendente que tiene múltiples pistas y objetos, que incluye una señal mezclada de manera descendente, se consigue de una manera que reduce el requisito de tasa de bits en comparación con la transmisión redundante o compresión sin pérdidas, mientras reduce los artefactos de mezcla ascendente. Se genera y transmite una señal residual comprimida junto con una mezcla total comprimida y al menos un objeto de audio comprimido. En el aspecto de recepción y mezcla ascendente la invención descomprime una señal mezclada de manera descendente y otros objetos comprimidos,

calcula una señal de mezcla ascendente aproximada y corrige señales de base específicas derivadas de la mezcla ascendente, restando una señal residual descomprimida. La invención por lo tanto permite que se use compresión con pérdidas en comparación con señales de audio mezcladas de manera descendente para transmisión a través de un canal de comunicación (o para almacenamiento). Tras la recepción y mezcla ascendente posterior, pueden recuperarse señales de base adicionales en sistemas aptos que proporcionan capacidad de múltiples objetos (mientras que los sistemas heredados pueden decodificar fácilmente una mezcla total sin mezcla ascendente). El método y aparato de la invención tiene tanto a) aspectos de compresión de audio y mezcla descendente, como b) un aspecto de descompresión de audio/mezcla ascendente, en el que la compresión debería entenderse que indica un método de reducción de tasa de bits (o reducción de tamaño de fichero), y en el que la mezcla descendente indica una reducción en recuento de canales u objetos, mientras que la mezcla ascendente indica un aumento en el recuento de canales recuperando y separando un canal u objeto previamente mezclado de manera descendente.

En su aspecto de descompresión y mezcla ascendente, en una realización preferida, al menos una de las representaciones comprimidas de C y de al menos una Bi se preparan por un método de compresión con pérdidas.

En su aspecto de compresión y mezcla descendente, se transmiten preferentemente (o de manera equivalente, se almacenan o graban) la señal de mezcla total comprimida E(C), el conjunto de (al menos una) señales de objeto comprimidas E(Bi), y la señal residual comprimida Ec(A).

En una realización del aspecto de compresión y mezcla descendente, la señal de referencia comprende la señal de mezcla de base A. En una realización alternativa, la señal de referencia es una aproximación de la señal de base A derivada comprimiendo la señal de base A por un método con pérdidas para formar una señal comprimida E(A), descomprimiendo a continuación la señal comprimida E(A) para obtener una señal de referencia (que es una aproximación de la señal de base A).

Como se usa en esta solicitud, a menos que el contexto lo exija claramente de otra manera, el término "conjunto" se usa para indicar un conjunto que tiene al menos un miembro, pero no requiere necesariamente que tenga una pluralidad de miembros. Este sentido se usa comúnmente en contextos matemáticos y no debería provocar ambigüedad. Estas y otras características y ventajas de la invención serán evidentes para los expertos en la materia a partir de la siguiente descripción detallada de realizaciones preferidas, tomadas junto con los dibujos adjuntos, en los que:

#### Breve descripción de los dibujos:

La Figura 1 es un diagrama de bloques de alto nivel que representa un sistema generalizado para comprimir y transmitir las señales compuestas que incluyen señales de audio mezcladas de una manera compatible hacia atrás, como es conocido en la técnica anterior;

La Figura 2 es un diagrama de flujo que muestra las etapas de un método para comprimir una señal de audio compuesta de acuerdo con una primera realización de la invención;

La Figura 3 es un diagrama de flujo que muestra las etapas de un método para descomprimir y mezclar de manera ascendente señales de audio, de acuerdo con un aspecto de descompresión de la invención;

La Figura 4 es un diagrama de flujo que muestra etapas de un método para comprimir una señal de audio compuesta de acuerdo con una realización alternativa de la invención;

La Figura 5 es un diagrama de bloques esquemático de un aparato para comprimir una señal de audio compuesta de acuerdo con una realización alternativa de la invención, consistente con el método de la Figura 2; y

La Figura 6 es un diagrama de bloques esquemático de un aparato para comprimir una señal de audio compuesta de acuerdo con una primera realización de la invención, consistente con el método de la Figura 4.

#### Descripción detallada de la invención:

Los métodos descritos en el presente documento se refieren a procesamiento de señales, y están dirigidos particularmente a procesamiento de señales de audio que representan sonido físico. Estas señales pueden representarse por señales electrónicas digitales. En el análisis, pueden mostrarse o analizarse formulaciones matemáticas continuas para ilustrar los conceptos; sin embargo, debería entenderse que algunas realizaciones operan en el contexto de una serie temporal de bytes o palabras digitales, formando dichos bytes o palabras una aproximación discreta de una señal analógica o (finalmente) un sonido físico. La señal digital discreta corresponde a una representación digital de una forma de onda de audio muestreada periódicamente. En una realización, puede usarse una tasa de muestreo de aproximadamente 48 mil muestras/segundo. Como alternativa pueden usarse tasas de muestreo superiores tales como 96 khz. El esquema de cuantización y resolución de bits puede elegirse para satisfacer los requisitos de una aplicación particular. Las técnicas y aparato descritos en el presente documento pueden aplicarse de manera interdependiente en un número de canales. Por ejemplo, pueden usarse en el contexto de un sistema de audio envolvente que tiene más de dos canales.

Como se usa en el presente documento, una "señal de audio digital" o "señal de audio" no describe una mera abstracción matemática, sino que, además de tener su significado habitual, indica información realizada en o llevada

por un medio físico no transitorio, apto para detección por una máquina o aparato. Esta expresión incluye señales grabadas o transmitidas, y no debería entenderse que incluye transporte por forma alguna de codificación, incluyendo modulación por codificación de pulsos (PCM), pero sin limitación a PCM. Las salidas o entradas, podrían codificarse o comprimirse por cualquiera de diversos métodos conocidos, incluyendo MPEG, ATRAC, AC3, o los métodos propietarios de DTS, Inc., como se describe en las Patentes de Estados Unidos 5.974.380; 5.978.762; y 6.487.535. Puede realizarse alguna modificación de los cálculos para adaptar esa compresión particular o método de codificación.

#### Vista general:

La Figura 1 muestra el entorno general con el que opera la invención, a un alto nivel de generalización. Como en la técnica anterior, un codificador 110 recibe una pluralidad de señales de audio independientes denominadas de manera arbitraria como A, B, mezcla de manera descendente dichas señales a una señal de mezcla total  $C (=A+B)$  con un mezclador 120, comprime las señales mezcladas de manera descendente con el compresor 130, a continuación transmite (o graba) las señales mezcladas de manera descendente de una manera que permitirá la reconstrucción de una aproximación razonable de las señales en un decodificador 160. Aunque únicamente se muestra la señal B en los dibujos (por simplicidad), la invención puede usarse con una pluralidad de señales u objetos independientes  $B_1, B_2, \dots, B_m$ . De manera similar, en la descripción que sigue hacemos referencia a un conjunto de objetos  $B_1, B_2, \dots, B_m$ ; debería entenderse que el conjunto de objetos consiste en al menos un objeto, es decir  $m \geq 1$ , sin limitación a un cierto número de objetos.

Además del codificador 110 y del decodificador 160, la Figura 1 muestra un canal de transmisión generalizado 150, que debería entenderse que incluye cualquier medio de transmisión o medio de grabación o almacenamiento, particularmente grabación en un medio de almacenamiento legible por máquina no transitorio. En el contexto de la invención, y en teoría de comunicación más en general, la grabación o almacenamiento combinados con posterior reproducción puede considerarse un caso especial de transmisión o comunicación de información, se entiende que la reproducción corresponde a recibir y decodificar la información codificada en general a un momento posterior y opcionalmente en una localización espacial diferente. Por lo tanto, el término "transmitir" puede indicar grabación en un medio de almacenamiento; "recibir" puede indicar lectura de un medio de almacenamiento; y "canal" puede incluir almacenamiento de información en un medio.

Es importante que las señales se transmitan a través del canal de transmisión en un formato multiplexado para mantener y conservar la relación síncrona entre las señales (A, B, C). El multiplexor y demultiplexor podría incluir combinaciones de métodos de empaquetamiento de bits y formateo de datos conocidos en la técnica. El canal de transmisión puede incluir también otras capas de codificación o procesamiento de información, tales como corrección de errores, comprobación de paridad u otras técnicas según sean apropiadas al canal o capas físicas como se describe en el modelo de capa OSI (por ejemplo).

Como se muestra, un decodificador recibe señales de audio comprimidas y mezcladas de manera descendente, demultiplexa dichas señales, descomprime dichas señales de una manera inventiva que permite la reconstrucción aceptable de una mezcla ascendente para reproducir una pluralidad de señales independientes (u objetos de audio). Las señales a continuación preferentemente se mezclan de manera ascendente para recuperar las señales originales (o una aproximación tan cerca como sea posible).

#### Teoría del funcionamiento:

Suponiendo que  $A, B_1, B_2, \dots, B_m$  son señales independientes (objetos), que se codifican en un *flujo de código* y se envían a un *representador*. El objeto distinguido A se denominará como el *objeto de base*, mientras que  $B = B_1, B_2, \dots, B_m$  se denominarán como *objetos regulares*. Hacemos referencia a un conjunto de objetos  $B_1, B_2, \dots, B_m$ ; pero debería entenderse el conjunto de objetos contiene al menos un objeto (es decir  $m \geq 1$ ), no limitado a un cierto número de objetos. En un sistema de audio basado en objetos, estamos interesados en representar objetos simultánea pero independientemente, de modo que, por ejemplo, cada objeto se representaría en una localización espacial diferente.

Para compatibilidad hacia atrás, requerimos que el flujo codificado sea interpretable por sistemas heredados que no son ni basados en objeto ni reconocen objetos. Tales sistemas pueden únicamente representar el *objeto* compuesto  $C = A + B_1 + B_2 + \dots + B_m$  de una versión codificada,  $E(C)$ , de C. Por lo tanto, requerimos que el flujo de código transmitido incluya  $E(C)$ , seguido por las descripciones de los objetos individuales, que se ignoran por los sistemas heredados. En métodos de la técnica anterior el flujo de código consistiría en  $E(C)$  seguido por las descripciones  $E(B_1), E(B_2), \dots, E(B_m)$  de los objetos regulares. El objeto de base A se recuperaría a continuación decodificando estas descripciones y estableciendo  $A = C - B_1 - B_2 - \dots - B_m$ . Debería observarse, sin embargo, que la mayoría de códecs de audio usados en la práctica son *con pérdidas*, que significa que la versión decodificada  $Q(X) = D(E(X))$  de un objeto codificado  $E(X)$  es únicamente una aproximación de X, y no necesariamente idéntica a ella. La precisión de la aproximación depende generalmente de la elección del códec  $\{E, D\}$  y del ancho de banda (o espacio de almacenamiento) disponible para el flujo de código.

Se deduce, por lo tanto, que cuando se usa un codificador con pérdidas, el decodificador no tendrá acceso a los objetos  $C, B_1, B_2, \dots, B_m$ , sino a versiones aproximadas  $Q(C), Q(B_1), Q(B_2), \dots, Q(B_m)$ , y únicamente podrá estimar  $A$  como

$$Q'(A) = Q(C) - Q(B_1) - Q(B_2) - \dots - Q(B_m).$$

Una aproximación de este tipo sufrirá de la acumulación de los errores en las codificaciones con pérdidas individuales. Esto a menudo dará como resultado, en la práctica, artefactos perceptivos objetables. En particular,  $Q'(A)$  puede ser una aproximación significativamente peor de  $A$  que  $Q(A)$ , y sus artefactos pueden estar estadísticamente correlacionados a los otros objetos, que no es el caso con  $Q(A)$ . En la práctica el residuo  $C - B_1 - B_2$  etc., estará correlacionado de manera audible a  $B_1 + B_2 + \dots$  (para compresión con pérdida). Nuestros oídos humanos pueden detectar correlaciones que son difíciles de detectar algorítmicamente.

De acuerdo con la invención, se evita alguna de la redundancia mencionada en relación con enfoques anteriores, mientras aún permite una reconstrucción aceptable de  $A$ . En lugar de incluir una (señal redundante)  $Q(A)$  en el flujo de código, incluimos una codificación  $E_c(\Delta)$ , donde  $\Delta$  es la señal residual:

$$\Delta = Q'(A) - A,$$

y  $E_c$  es un codificador con pérdidas para  $\Delta$  (no necesariamente el mismo que  $E$ ).  $D_c$  es un decodificador para  $E_c$ , y

$$R(\Delta) = D_c(E_c(\Delta)).$$

En el lado del decodificador, una aproximación de  $A$  se obtiene como

$$Q_c(A) = Q'(A) - R(\Delta).$$

#### Método de la primera realización:

##### 1. Codificador

El método de codificación descrito matemáticamente anteriormente puede describirse de manera procedural como una secuencia de acciones, como se muestra en la **Figura 2**. Como se ha descrito anteriormente, al menos un objeto distinguido  $A$  se denominará como el *objeto de base*, mientras que  $B_1, B_2, \dots, B_m$  se denominarán como *objetos regulares*. Por brevedad, hacemos referencia a los objetos regulares de manera colectiva como  $B$  a continuación, se entiende que el conjunto de todos (al menos uno) objetos regulares  $B_1, B_2, \dots, B_m$  puede designarse como  $\{B_i\}$ ; en contraste,  $B = B_1 + B_2 + \dots + B_m$  indica la mezcla del objeto regular  $B_1, B_2, \dots, B_m$ . El método comienza con una señal mezclada  $C = A + B$ . Será evidente que la mezcla de  $A + B$  podría hacerse como una etapa preliminar, o las señales podrían proporcionarse según se han mezclado previamente. La señal  $A$  también es necesaria; puede recibirse de manera separada o reconstruirse por la resta de  $B$  de  $C$ . El conjunto de (al menos uno) objetos regulares  $\{B_i\}$  también se requiere y se usa por el codificador como se describe a continuación.

En primer lugar, el codificador comprime (etapa 210) las señales  $A, \{B_i\}$  y  $C$  de manera separada usando un método de codificación con pérdidas para obtener señales comprimidas correspondientes indicadas  $E(A), \{E(B_i)\}$ , y  $E(C)$  respectivamente. (La notación  $\{E(B_i)\}$  indica el conjunto de objetos codificados que cada uno corresponde con un respectivo objeto regional que pertenece al conjunto de señales  $\{B_i\}$ , cada señal de objeto codificada individualmente por  $E$ ). El codificador a continuación descomprime (etapa 220)  $E(C)$  y  $\{E(B_i)\}$  por un método complementario al usado para comprimir  $C$  y  $\{B_i\}$ , para producir las señales reconstruidas  $Q(C)$  y  $\{Q(B_i)\}$ . Estas señales aproximan la original  $C$  y  $\{B_i\}$  (diferiendo puesto que se comprimieron y a continuación descomprimieron usando un método de compresión/descompresión con pérdidas).  $\{Q(B_i)\}$  se resta a continuación de  $Q(C)$  por la etapa de mezcla sustractiva 230 para producir una señal de mezcla ascendente modificada  $Q'(A)$ , que es una aproximación de la original  $A$  que difiere de  $A$  por errores introducidos en codificación con pérdidas seguido por la mezcla. A continuación, la señal  $A$  (una señal de referencia) se resta de la señal de mezcla ascendente modificada  $Q'(A)$  en una segunda etapa de mezcla 240 para obtener una señal residual  $\Delta = Q'(A) - A$  (etapa 130). La señal residual  $\Delta$  se comprime a continuación (etapa 250) por un método de compresión que designamos como  $E_c$ , donde  $E_c$  no es necesariamente el mismo método o dispositivo de compresión que  $E$  (usado en la etapa 210 para comprimir las señales  $A, \{B_i\}$ , o  $C$ ). Preferentemente, para reducir los requisitos de ancho de banda  $E_c$  debería ser un codificador con pérdidas para  $\Delta$  elegido para adaptar las características de  $\Delta$ . Sin embargo, en una realización alternativa menos optimizada para ancho de banda,  $E_c$  podría ser un método de compresión sin pérdidas.

Obsérvese que el método anteriormente descrito requiere las etapas 210 y 220 de compresión y descompresión sucesivas (según se aplican a las señales  $\{B_i\}$  y  $C$ ). En estas etapas, y en el método alternativo descrito a continuación, la complejidad de cálculo y el tiempo pueden reducirse en algunos casos únicamente realizando las porciones con pérdidas de la compresión (y descompresión). Por ejemplo, muchos métodos de descompresión con pérdidas tales como el códec DTS descrito en la Patente de Estados Unidos 5974380 requieren aplicaciones sucesivas de ambas etapas con pérdidas (filtración en subbandas, asignación de bits, re-cuantificación en

subbandas) seguido por etapas sin pérdidas (aplicación de un libro de códigos, reducción de entropía). En tal caso es suficiente omitir las etapas sin pérdidas tanto en la codificación como en la decodificación, simplemente realizando las etapas con pérdidas. La señal reconstruida mostraría aún todos los efectos de transmisión con pérdidas, pero se ahorran muchas etapas computacionales.

5 El codificador a continuación transmite (etapa 260)  $R = E_c(\Delta)$ ,  $E(C)$  y  $\{E(B_i)\}$ . Preferentemente el método de codificación también incluye la etapa opcional de multiplexación o reformato de las tres señales en un paquete multiplexado para su transmisión o grabación. Cualquiera de los métodos conocidos de multiplexación podría usarse, con la condición de que se usen algunos medios para conservar o reconstruir la sincronización temporal de las tres señales separadas o relacionadas. Debería tenerse en cuenta que el esquema de cuantificación diferente puede usarse para todas las tres señales, y que el ancho de banda puede distribuirse entre las señales. Cualquiera de los muchos métodos conocidos de compresión de audio con pérdidas podría usarse para  $E$ , incluyendo MP3, AAC, WMA o DTS (por nombrar solo unos pocos).

15 Este enfoque ofrece al menos las siguientes ventajas: en primer lugar, la señal de "error"  $\Delta$  se espera que sea de menor potencia y entropía que los objetos originales. Teniendo potencia reducida en comparación con  $A$ , la señal de error  $\Delta$  puede codificarse con menos bits que el objeto  $A$  al que ayuda a reconstruir. Por lo tanto, el enfoque propuesto se espera que sea más económico que el método de descripción redundante analizado anteriormente (en la sección de Antecedentes). En segundo lugar, el codificador  $E$  puede ser cualquier codificador de audio (por ejemplo, MP3, AAC, WMA, etc.), y obsérvese especialmente que el codificador puede ser, y en realizaciones preferidas es, un codificador con pérdidas que emplea principios psicoacústicos. (El correspondiente decodificador sería también por supuesto un decodificador con pérdidas correspondiente). En tercer lugar, el codificador  $E_c$  no necesita ser un codificador de audio convencional, y puede estar optimizado para la señal  $\Delta$ , que no es una señal de audio convencional. De hecho, las consideraciones perceptuales en el diseño y optimización de  $E_c$  serán diferentes de aquellas en el diseño de un códec de audio convencional. Por ejemplo, los códecs de audio perceptuales no siempre buscan maximizar la SNR en todas las partes de la señal; en su lugar, en ocasiones se busca un régimen de SNR instantánea más "constante", donde se permiten errores mayores cuando la señal es más intensa. De hecho, esto es una fuente principal de los artefactos que resultan de  $B_i$  que se hallan en  $Q'(A)$ . Con  $E_c$ , buscamos eliminar estos artefactos tanto como sea posible, por lo que una maximización de SNR instantánea directa parece más apropiada en este caso.

El método de decodificación de acuerdo con la invención se muestra en la **Figura 3**. Como una etapa opcional preliminar 300, el decodificador debe recibir y demultiplexar el flujo de datos para recuperar  $E_c(\Delta)$ ,  $\{E(B_i)\}$  y  $E(C)$ . En primer lugar, (etapa 310) el decodificador recibe los flujos de datos comprimidos (o ficheros)  $E_c(\Delta)$ ,  $\{E(B_i)\}$  y  $E(C)$ . A continuación el decodificador descomprimirá (etapa 320) cada uno de los flujos de datos (o ficheros)  $E_c(\Delta)$ ,  $\{E(B_i)\}$  y  $E(C)$  para obtener representaciones reconstruidas  $\{Q(B_i)\}$ ,  $Q(C)$  y  $R_c(\Delta) = D_c(E_c(\Delta))$  donde  $D_c$  es el método de descompresión inverso al método de compresión  $E_c$ , y donde los métodos de descompresión para  $\{E(B_i)\}$  y  $E(C)$  son aquellos complementarios a los métodos de compresión usados para  $\{B_i\}$  y  $C$ . Las señales  $Q(C)$  y  $\{Q(B_i)\}$  se mezclan de manera sustractiva (etapa 330) para recuperar  $Q'(A) = Q(C) - \sum Q(B_i)$ . Esta señal  $Q'(A)$  es una aproximación de  $A$  que difiere de la original  $A$  puesto que se reconstruyó de una mezcla sustractiva de  $Q(C)$  y  $\{Q(B_i)\}$ , ambas de las cuales se transmitieron por métodos de códec con pérdidas. En el método de decodificación y mezcla ascendente de la invención, la señal de aproximación  $Q'(A)$  se mejora a continuación restando (etapa 340) el residuo reconstruido  $R(\Delta)$  para obtener  $Q_c(A) = Q'(A) - R(\Delta)$ . Las señales de réplica recuperadas  $Q_c(A)$ ,  $Q(C)$ ,  $\{Q(B_i)\}$  pueden a continuación reproducirse o emitirse para reproducción (etapa 350) como una mezcla ascendente ( $A$ ,  $\{B_i\}$ ). La señal de mezcla descendente  $Q(C)$  también está disponible para su emisión para sistemas que tienen menos canales (o como una elección basándose en el control o preferencia del consumidor).

Se apreciará que el método de la invención requiere transmisión de algunos datos redundantes. Sin embargo, el tamaño de fichero (o requisito de tasa de bits) para el método de la invención es menor que el requerido para cualquiera de a) uso de codificación sin pérdidas para todos los canales, o b) transmitir una descripción redundante de objetos codificados con pérdidas más mezcla ascendente codificada con pérdidas. En un experimento, el método de la invención se usó para transmitir una mezcla ascendente  $A+B$  (para un único objeto  $B$ ), junto con el canal de base  $A$ . Los resultados se muestran en la Tabla 1. Puede observarse que el método de descripción redundante (técnica anterior) requeriría 309 KB para transmitir la mezcla; en contraste, el método de la invención requeriría únicamente 251 KB para la misma información (más alguna tara mínima para los campos de multiplexación y encabezamiento). Este experimento no representa los límites de mejora que pueden obtenerse optimizando adicionalmente los métodos de compresión.

En una realización alternativa del método, como se muestra en la **Figura 4**, el método de codificación difiere en que la señal residual  $\Delta$  se deriva de la diferencia entre  $Q'(A) = D(E(C)) - \sum D(E(B_i))$  y  $Q(A)$  (en lugar de  $A$ ). Esta realización es particularmente apropiada en una aplicación en la que se desea la reconstrucción de  $A$  y se espera que alcance aproximadamente la misma calidad que la reconstrucción de  $B$  y  $C$  (no hay necesidad de esforzarse por una reconstrucción de fidelidad superior de  $A$ ). Este a menudo es el caso en un sistema de entretenimiento de audio.

65 Obsérvese que en la realización alternativa,  $Q'(A)$  es la señal reproducida tomando la diferencia entre a) la versión codificada después decodificada de la mezcla descendente de  $C$ , y b) los objetos de base reconstruidos  $\{Q(B_i)\}$

reproducidos decodificando la mezcla de base codificada con pérdidas B.

Haciendo referencia ahora a la Figura 4, en la alternativa del método, el codificador comprime (etapa 410) las señales A, {Bi}, y C de manera separada usando un método de codificación con pérdidas para obtener tres señales comprimidas correspondientes indicadas EA, {E(Bi)} y E(C) respectivamente. El codificador a continuación descomprime E(A) (etapa 420) por un método complementario al usado para comprimir A que produce Q(A) que es una aproximación de A (que difiere puesto que se comprimió y a continuación descomprimió usando un método de compresión/descompresión con pérdidas). El método alternativo a continuación descomprime (etapa 430) tanto E(C) como {E(Bi)} mediante respectivos métodos complementarios a aquellos usados para codificar C y {Bi}. Las señales reconstruidas resultantes Q(C) y {Q(Bi)} son aproximaciones a la original {Bi} y C, difieren debido a las imperfecciones introducidas por los métodos de codificación y decodificación con pérdidas. El método alternativo a continuación en la etapa 440 resta  $\Sigma Q(Bi)$  de Q(C) para obtener la señal de diferencia Q'(A). Q'(A) es otra aproximación de A, que difiere debido a que la compresión con pérdidas se usó en la mezcla descendente transmitida. Se obtiene una señal residual  $\Delta$  (etapa 450) restando Q(A) de Q'(A).

La señal residual  $\Delta$  se comprime a continuación en la etapa 460 por el método de codificación Ec (que podría diferir de E). Como en la primera realización anteriormente descrita, Ec es preferentemente un códec con pérdidas adecuado a las características de la señal residual. El codificador a continuación transmite (etapa 470)  $R=Ec(\Delta)$ , E(C) y {E(Bi)} a través de un canal de transmisión con la relación síncrona conservada. Preferentemente el método de codificación también incluye multiplexación o reformato de las tres señales en un paquete multiplexado para transmisión o grabación. Cualquiera de los métodos conocidos de multiplexación podría usarse, con la condición de que se usen algunos medios para conservar o reconstruir la sincronización temporal de las tres señales separadas o relacionadas. Debería tenerse en cuenta que puede usarse el esquema de cuantificación diferente para todas las tres señales, y que el ancho de banda puede distribuirse entre las señales. Cualquiera de los muchos métodos conocidos de compresión de audio podría usarse para E, incluyendo MP3, AAC, WMA o DTS (por nombrar únicamente unos pocos).

Las señales codificadas por el método de codificación alternativo pueden decodificarse por el mismo método de decodificación anteriormente descrito en relación con la Figura 3. El decodificador restará la señal residual reconstruida para mejorar la aproximación de la señal de mezcla ascendente, Q(A), reduciendo de esta manera la diferencia entre la señal de réplica reconstruida Q(A) y la señal original A. Ambas realizaciones de la invención están unidas por la generalidad de que generan en el codificador una señal residual o de error  $\Delta$  que representa la diferencia que se espera después de la decodificación y mezcla ascendente de una señal para extraer un objeto con privilegio A. La señal de error  $\Delta$  se comprime y transmite en ambas realizaciones (o de manera equivalente, se recupera o almacena). En ambas realizaciones el decodificador descomprime la señal de error comprimida  $\Delta$  y la resta de la señal de mezcla ascendente reconstruida que aproxima el objeto con privilegio A.

El método de la realización alternativa puede tener algunas ventajas perceptuales en ciertas aplicaciones. Cuáles de las alternativas se prefieren en la práctica puede depender de los parámetros específicos del sistema y los objetivos de optimización específicos.

En otro aspecto, la invención incluye un aparato para comprimir o codificar señales de audio mezcladas como se muestra en la Figura 5. En una primera realización del aparato, se proporcionan las señales C (=A+B mezcla de objeto) y B en la entrada 510 y 512, respectivamente. La señal C se codifica por el codificador 520 para producir la señal codificada E(C); las señales {Bi} se codifican por el codificador 530 para producir la segunda señal codificada {E(Bi)}. E(C) y {E(Bi)} se decodifican a continuación por los decodificadores 540 y 550, respectivamente, para producir las señales reconstruidas Q(C) y {Q(Bi)}. Las señales reconstruidas Q(C) y {Q(Bi)} se mezclan de manera sustractiva en el mezclador 560 para producir la señal de diferencia Q'(A). Esta señal de diferencia difiere de la señal original A en que se obtiene mediante la mezcla a partir de una mezcla total reconstruida Q(C) y los objetos reconstruidos {Q(Bi)}; se introducen artefactos o errores tanto debido a que el codificador 520 es un codificador con pérdidas, como debido a que la señal se deriva por resta (en el mezclador 560). La señal reconstruida Q'(A) se resta a continuación de la señal A (introducida en 570) y la diferencia  $\Delta$  se comprime por un segundo codificador 580 que en una realización preferida opera por un método diferente que el compresor 520 para producir una señal residual comprimida Ec( $\Delta$ ).

En una realización alternativa del aparato codificador, mostrada en la Figura 6, se proporcionan las señales C (=A+B mezcla de objeto) y B en la entrada 510 y 512, respectivamente. La señal C se codifica por el codificador 520 para producir la señal codificada E(C); las señales {Bi} se codifican por el codificador 530 para producir la segunda señal codificada E(B). E(C) y {E(Bi)} se decodifican a continuación por los decodificadores 540 y 550, respectivamente, para producir las señales reconstruidas Q(C) y {Q(Bi)}. Las señales reconstruidas Q(C) y Q(B) se mezclan de manera sustractiva en el mezclador 560 para producir la señal de diferencia Q'(A). Esta señal de diferencia difiere de la señal original A en que se obtiene mediante la mezcla a partir de una mezcla total reconstruida Q(C) y los objetos reconstruidos {Q(Bi)}; se introducen artefactos o errores tanto debido a que el codificador 520 es un codificador con pérdidas, como debido a que la señal se deriva por resta (en el mezclador 560). Hasta ahora la realización alternativa se asemeja a la primera realización.



En la realización alternativa del aparato, la señal A recibida en la entrada 570 se codifica por el codificador 572 (que puede ser el mismo u operar por los mismos principios que los codificadores con pérdidas 520 y 530) a continuación la salida codificada de 572 se decodifica de nuevo por un decodificador complementario 574 para producir una aproximación reconstruida Q(A) que se diferencia de A debido a la naturaleza con pérdidas del codificador 572. La

5 señal reconstruida Q(A) se resta a continuación de Q'(A) en el mezclador 560, y la señal residual resultante se codifica por un segundo codificador 580 (método diferente del usado en los codificadores con pérdidas 520 y 530). Las salidas E(C), {E(Bi)} y E( $\Delta$ ) se hacen disponibles a continuación para su transmisión o grabación, preferentemente en algún formato multiplexado o cualquier otro método que permita sincronización.

10 Será evidente que el contenido codificado por el primero o métodos alternativos o por el aparato de codificación (Figura 6) puede decodificarse por el decodificador de la Figura 3. El decodificador requiere una señal de error comprimida, pero no necesita ser sensible a la manera en la que se calcula el error. Esto deja oportunidad para mejora futura en el códec sin cambiar el diseño de decodificador.

15 Los métodos descritos en el presente documento pueden implementarse en un dispositivo de electrónica de consumidor, tal como un ordenador de fin general, estación de trabajo de audio digital, reproductor de DVD o BD, sintonizador de TV, reproductor de CD, reproductor portátil, dispositivo de audio/vídeo de Internet, una consola de juegos, un teléfono móvil, auriculares o similares. Un dispositivo electrónico de consumidor puede incluir una Unidad de Procesamiento Central (CPU), que puede representar uno o más tipos de procesadores, tales como un IBM  
20 PowerPC, Intel Pentium (x86) procesadores, y así sucesivamente. Una Memoria de Acceso Aleatorio (RAM) almacena temporalmente resultados de las operaciones de procesamiento de datos realizadas por la CPU, y puede estar interconectada a la misma normalmente mediante un canal de memoria especializada. El dispositivo electrónico de consumidor puede incluir también dispositivos de almacenamiento permanente tal como un disco duro, que puede también estar en comunicación con la CPU a través de un bus de E/S. Otros tipos de dispositivos  
25 de almacenamiento tales como unidades de cinta o unidades de discos ópticos pueden estar también conectados. Una tarjeta de gráficos puede también estar conectada a la CPU mediante un bus de vídeo, y transmitir señales representativas de datos de visualización al monitor de visualización. Los dispositivos de entrada de datos periféricos externos, tales como un teclado o un ratón, pueden estar conectados al sistema de reproducción de audio a través de un puerto USB. Un controlador de USB puede traducir datos e instrucciones a y desde la CPU para periféricos  
30 externos conectados al puerto USB. Dispositivos adicionales tales como impresoras, micrófonos, altavoces, auriculares y similares pueden estar conectados al dispositivo electrónico de consumidor.

El dispositivo electrónico de consumidor puede utilizar un sistema operativo que tiene una interfaz de usuario gráfica (GUI), tal como WINDOWS de Microsoft Corporation en Redmond, Washington, MAC OS de Apple, Inc. en  
35 Cupertino, CA, diversas versiones de GUI móviles diseñadas para sistemas operativos móviles tales como Android, y así sucesivamente. El dispositivo electrónico de consumidor puede ejecutar uno o más programas informáticos. En general, el sistema operativo y los programas informáticos se realizan de manera tangible en un medio legible por ordenador no transitorio, por ejemplo uno o más de los dispositivos de almacenamiento de datos fijos y/o extraíbles incluyendo el disco duro. Tanto el sistema operativo como los programas informáticos pueden cargarse a partir de  
40 los dispositivos de almacenamiento de datos anteriormente mencionados en la RAM para su ejecución por la CPU. Los programas informáticos pueden comprender instrucciones que, cuando se leen y ejecutan por la CPU, provocan que la misma realice las etapas para ejecutar las etapas o características de las realizaciones descritas en el presente documento.

45 Las realizaciones descritas en el presente documento pueden tener muchas configuraciones y arquitecturas diferentes. Cualquiera de tal configuración o arquitectura puede sustituirse fácilmente. Un experto en la materia reconocerá las secuencias anteriormente descritas que son las más comúnmente utilizadas en medios legibles por ordenador, pero hay otras secuencias existentes que pueden sustituirse.

50 Los elementos de una realización pueden implementarse por hardware, firmware, software o cualquier combinación de los mismos. Cuando se implementan como hardware, las realizaciones descritas en el presente documento pueden emplearse en un procesador de señales de audio o distribuirse entre diversos componentes de procesamiento. Cuando se implementan en software, los elementos de una realización pueden incluir los segmentos de código para realizar las tareas necesarias. El software puede incluir el código real para llevar a cabo las  
55 operaciones descritas en una realización o código que emula o simula las operaciones. El programa o segmentos de código pueden almacenarse en un procesador o medio accesible por máquina o transmitirse por una señal de datos de ordenador realizada en una onda portadora, o una señal modulada por una portadora, a través de un medio de transmisión. El medio legible o accesible por procesador o el medio legible o accesible por máquina pueden incluir cualquier medio que pueda almacenar, transmitir o transferir información. En contraste, un medio de  
60 almacenamiento legible por ordenador o almacenamiento informático no transitorio puede incluir un dispositivo de almacenamiento de máquina informática física pero que no abarca una señal.

Ejemplos del medio legible por procesador incluyen un circuito electrónico, un dispositivo de memoria de semiconductores, una memoria de solo lectura (ROM), una memoria flash, una ROM borrable (EROM), un disquete flexible, un disco compacto (CD) ROM, un disco óptico, un disco duro, un medio de fibra óptica, un enlace de frecuencia de radio (RF), etc. La señal de datos informática puede incluir cualquier señal que pueda propagarse a

través de un medio de transmisión tal como canales de red electrónicos, fibras ópticas, aire, electromagnéticos, enlaces de RF, etc. Los segmentos de código pueden descargarse mediante redes informáticas tales como Internet, Intranet, etc. El medio accesible por máquina puede realizarse en un artículo de fabricación. El medio accesible por máquina puede incluir datos que, cuando se acceden por una máquina, provocan que la máquina realice la operación descrita a continuación. El término "datos", además de tener su significado habitual, en este punto hace referencia a cualquier tipo de información que se codifica para fines legibles por máquina. Por lo tanto, puede incluir programa, código, un fichero, etc.

Toda o parte de diversas realizaciones pueden implementarse por software que se ejecuta en una máquina, tal como un procesador de hardware que comprende circuitería lógica digital. El software puede tener varios módulos acoplados entre sí. El procesador de hardware podría ser un microprocesador digital programable, o procesador de señales digitales (DSP) programable especializado, un campo de matriz de puertas programables, un ASIC, u otro procesador digital. En una realización, por ejemplo, todas las etapas de un método de acuerdo con la invención (ya sea en aspecto de codificador o aspecto de decodificador) podrían llevarse a cabo de manera adecuada por uno o más ordenadores digitales programables que ejecutan todas las etapas de manera secuencial bajo el control de software. Un módulo de software puede estar acoplado a otro módulo para recibir variables, parámetros, argumentos, punteros, etc., y/o para generar o pasar resultados, variables actualizadas, punteros, etc. Un módulo de software puede ser también un controlador o interfaz de software para interactuar con el sistema operativo que se ejecuta en la plataforma. Un módulo de software puede incluir también un controlador de hardware para configurar, establecer, inicializar, enviar o recibir datos a y desde un dispositivo de hardware.

Pueden describirse diversas realizaciones como uno o más procesos, que pueden representarse como un gráfico de flujo, un diagrama de flujo, un diagrama de estructura, o un diagrama de bloques. Aunque un diagrama de bloques puede describir las operaciones como un proceso secuencial, muchas de las operaciones pueden realizarse en paralelo o de manera concurrente. Además, el orden de las operaciones puede reorganizarse. Un proceso se termina cuando sus operaciones están completadas. Un proceso puede corresponder a un método, un programa, un procedimiento, o similares.

A través de toda esta solicitud, se ha hecho referencia de manera frecuente a señales de adición, sustracción o "mezcla de manera sustractiva". Se reconocerá fácilmente que las señales pueden mezclarse de diversas maneras con resultados equivalentes. Por ejemplo, para restar una señal arbitraria F para G ( $G-F$ ), puede restarse directamente usando entradas diferenciales, o puede invertirse de manera equivalente una de las señales y a continuación añadir (ejemplo:  $G + (-F)$ ). Pueden concebirse otras operaciones equivalentes, incluyendo algunas la introducción de desplazamientos de fase. Términos tales como "restar" o "mezcla de manera sustractiva" se pretende que abarquen tales variaciones equivalentes. De manera similar, son posibles métodos variantes de adición de señal y se contemplan como "mezcla".

Aunque se han mostrado y descrito varias realizaciones ilustrativas de la invención, se les ocurrirán numerosas variaciones y realizaciones alternativas a los expertos en la materia. Tales variaciones y realizaciones alternativas se contemplan y pueden realizarse sin alejarse del alcance de la invención como se define en las reivindicaciones adjuntas.

## REIVINDICACIONES

1. Un método de descompresión y mezcla ascendente de una señal de audio compuesta comprimida y mezclada de manera descendente, que comprende las etapas:

5 recibir una representación comprimida  $E(C)$  de una señal de mezcla total  $C$ , una representación comprimida  $E_c(\Delta)$  de una señal residual  $\Delta$ , y un conjunto de representaciones comprimidas  $\{E(B_i)\}$  de respectivas señales de objeto de audio  $\{B_i\}$ , en donde la representación comprimida  $E(C)$  de la señal de mezcla total  $C$  es la representación comprimida  $E(C)$  de una señal de mezcla total  $C$  que comprende una señal de base  $A$  que se ha  
10 mezclado con un conjunto de señales de objeto de audio  $\{B_i\}$ ; en donde el conjunto de representaciones comprimidas  $\{E(B_i)\}$  de las señales de objeto de audio  $\{B_i\}$  incluye al menos una representación comprimida de una correspondiente señal de objeto  $B_i$ ; descomprimir la representación comprimida  $E(C)$  de la señal de mezcla total  $C$  para obtener una señal de mezcla total aproximada  $Q(C)$ ;  
15 descomprimir la representación comprimida  $E_c(\Delta)$  de la señal residual  $\Delta$  para obtener una señal residual reconstruida  $R_c(\Delta)$ ; descomprimir el conjunto de representaciones comprimidas  $\{E(B_i)\}$  de señales de objeto de audio  $\{B_i\}$  para obtener un conjunto de señales de objeto reconstruidas  $\{Q(B_i)\}$ , teniendo dicho conjunto una o más señales de objeto reconstruidas  $Q(B_i)$  como miembros; mezclar de manera sustractiva la señal de mezcla total aproximada  $Q(C)$  y el conjunto completo de señales de objeto reconstruidas  $\{Q(B_i)\}$  para obtener una primera aproximación  $Q'(A)$  de la señal de base  $A$ ; y  
20 mezclar de manera sustractiva la señal residual reconstruida  $R_c(\Delta)$  con la primera aproximación  $Q'(A)$  de la señal de base  $A$ , para obtener una aproximación mejorada  $Q_c(A)$  de la señal de base.

25 2. El método de la reivindicación 1, en el que dicho conjunto de representaciones comprimidas  $\{E(B_i)\}$  de las señales de objeto de audio comprende una representación comprimida de una señal de objeto de audio correspondiente.

3. El método de la reivindicación 1, en el que al menos una de las representaciones comprimidas  $E(C)$ ,  $\{E(B_i)\}$ ,  $E_c(\Delta)$  se prepara por un método de compresión con pérdidas.

30 4. El método de la reivindicación 3, en el que la representación comprimida  $E_c(\Delta)$  de la señal residual  $\Delta$  se prepara: mezclando de manera sustractiva una señal de referencia  $R$  con una aproximación reconstruida  $Q'(A)$  de la señal de base  $A$  para obtener una señal residual  $\Delta$  que representa la diferencia; y comprimiendo la señal residual  $\Delta$ .

35 5. El método de la reivindicación 1, que comprende adicionalmente: hacer que al menos una de la señal de base corregida  $Q'(A)$ , las señales de objeto reconstruidas  $\{Q(B_i)\}$  y la señal de mezcla total aproximada  $Q(C)$  se reproduzcan como un sonido.

40 6. El método de la reivindicación 1, en el que la etapa de descompresión del conjunto de representaciones comprimidas  $\{E(B_i)\}$  de respectivas señales de objeto de audio  $\{B_i\}$  comprime la descompresión de una pluralidad representaciones comprimidas para obtener una respectiva pluralidad de señales de objeto reconstruidas  $\{Q(B_i)\}$ ; y en donde dicha etapa de mezcla de manera sustractiva de la señal de mezcla total aproximada  $Q(C)$  y el conjunto completo de señales de objeto reconstruidas  $\{Q(B_i)\}$  incluye restar de  $Q(C)$  la pluralidad completa de señales de  
45 objeto reconstruidas  $\{Q(B_i)\}$ , para obtener la primera aproximación  $Q'(A)$  de la señal de base  $A$ .

7. El método de la reivindicación 6, en el que la representación comprimida  $E_c(\Delta)$  de la señal residual  $\Delta$  se prepara:

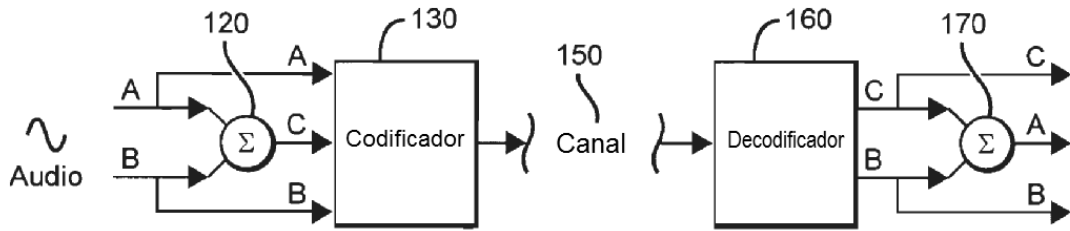
50 mezclando de manera sustractiva una señal de referencia  $R$  con la primera aproximación  $Q'(A)$  de la señal de base  $A$  para obtener una señal residual  $\Delta$  que representa la diferencia; y comprimiendo la señal residual  $\Delta$ .

8. Un método de compresión de una señal de audio compuesta que comprende una señal de mezcla total  $C$ , un conjunto de al menos una de las señales de objeto de audio  $\{B_i\}$  y una señal de base  $A$ , en donde dicha señal de  
55 mezcla total  $C$  comprende una señal de base  $A$  mezclada con el conjunto de señales de objeto de audio  $\{B_i\}$ , teniendo dicho conjunto de señales de objeto de audio  $\{B_i\}$  al menos una señal de objeto de miembro  $B_i$ , comprendiendo el método las etapas:

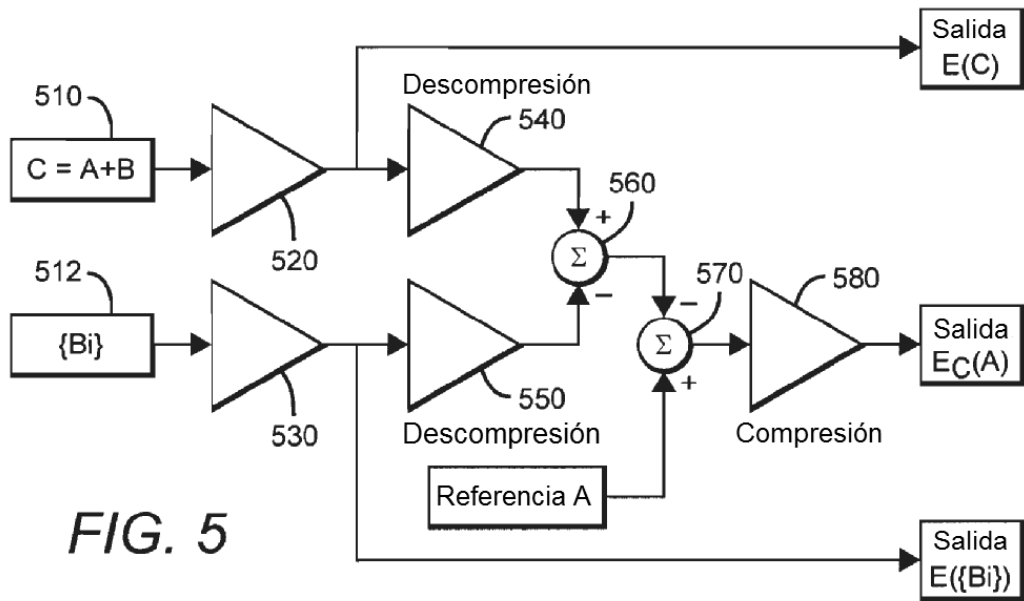
60 comprimir la señal de mezcla total  $C$  y el conjunto completo de señales de objeto de audio  $\{B_i\}$  por un método de compresión con pérdidas, para producir la señal de mezcla total comprimida  $E(C)$  y un conjunto comprimido de señales de objeto  $E(\{B_i\})$ , respectivamente; descomprimir la señal de mezcla total comprimida  $E(C)$  y el conjunto de señales de objeto comprimidas  $E(\{B_i\})$  para obtener una reconstruida  $Q(C)$  y un conjunto reconstruido de al menos unas señales de objeto  $Q(\{B_i\})$ ; mezclar de manera sustractiva la señal reconstruida  $Q(C)$  y una mezcla completa del conjunto de señales  
65 reconstruidas  $Q(\{B_i\})$  para producir una señal de base aproximada  $Q'(A)$ ; restar una señal de referencia de dicha señal de base aproximada  $Q'(A)$  para producir una señal residual  $\Delta$ ; y

comprimir la señal residual  $\Delta$  para obtener una señal residual comprimida  $E_c(\Delta)$ .

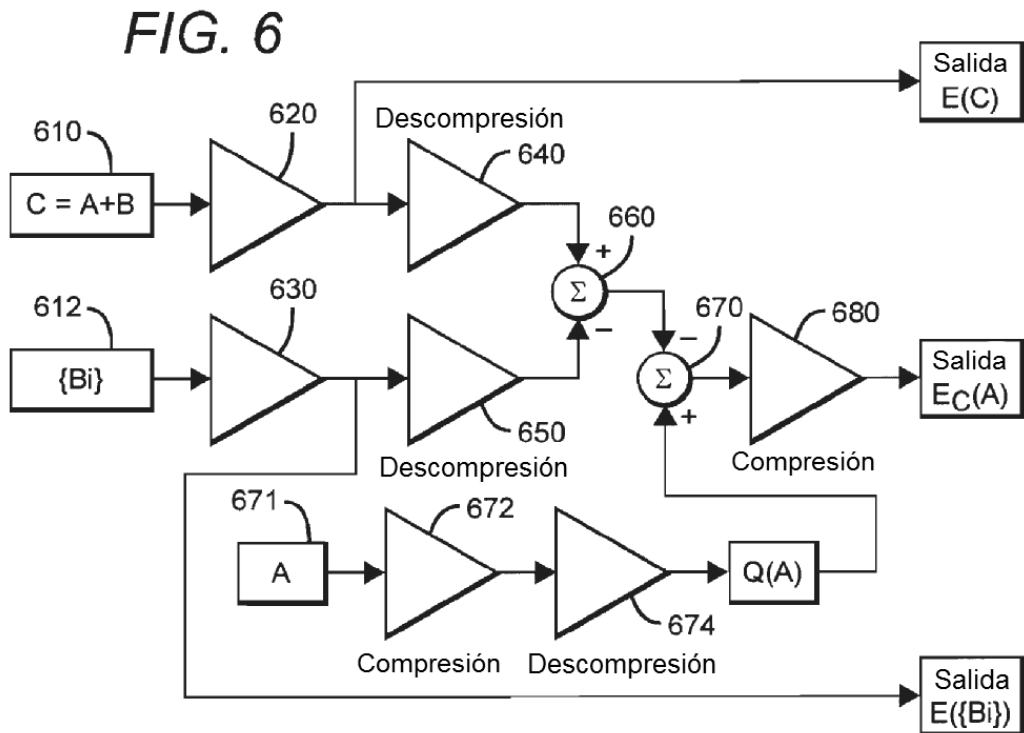
9. El método de la reivindicación 8, en el que dicho conjunto de al menos unas señales objeto  $\{B_i\}$  comprende únicamente una señal de objeto.
- 5 10. El método de la reivindicación 9, que comprende adicionalmente la etapa:  
transmitir una señal compuesta que comprende la señal de mezcla total comprimida  $E(C)$ , la señal de objeto comprimida  $E(\{B_i\})$  y la señal residual comprimida  $E(\Delta)$ .
- 10 11. El método de la reivindicación 9, en el que dicha señal de referencia comprende la señal de base A.
12. El método de la reivindicación 9, en el que dicha etapa de compresión de la señal residual comprende comprimir la señal residual por un método diferente de un método usado para comprimir la señal de mezcla total C.
- 15 13. El método de la reivindicación 8, en el que dicho conjunto de al menos una señal de objeto  $\{B_i\}$  comprime una pluralidad de señales de objeto.
14. El método de la reivindicación 13, en el que dicha señal de referencia comprende la señal de base A.
- 20 15. El método de la reivindicación 13, en el que dicha etapa de compresión de la señal residual comprende comprimir la señal residual por un método diferente de un método usado para comprimir la señal de mezcla total C.



**FIG. 1**  
TÉCNICA ANTERIOR



**FIG. 5**



**FIG. 6**

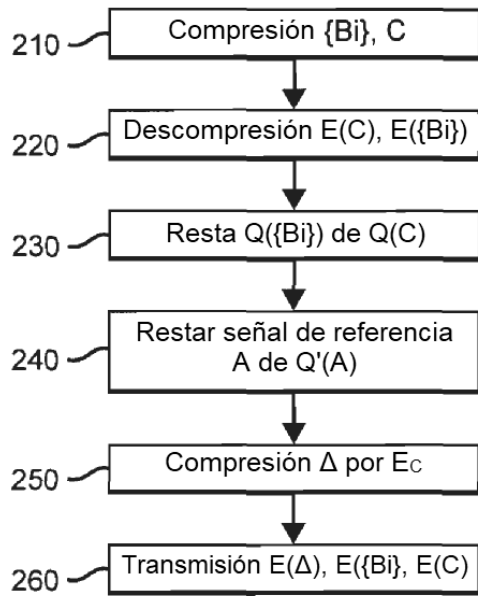


FIG. 2

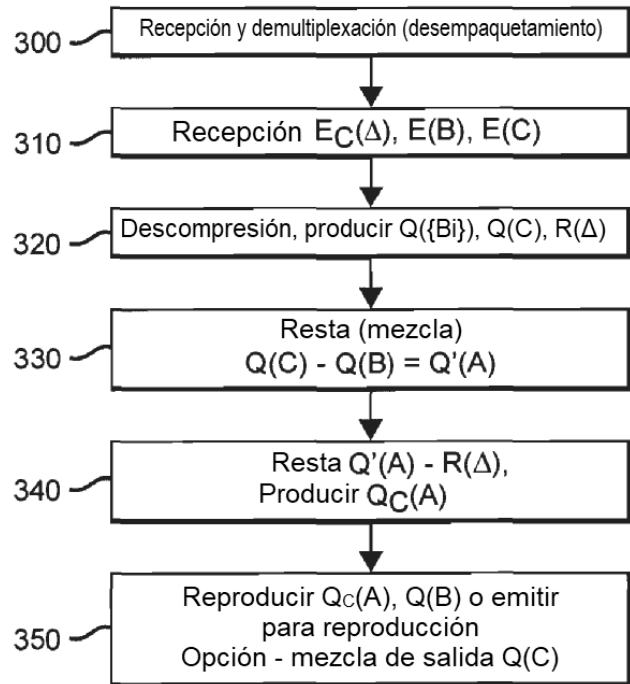


FIG. 3

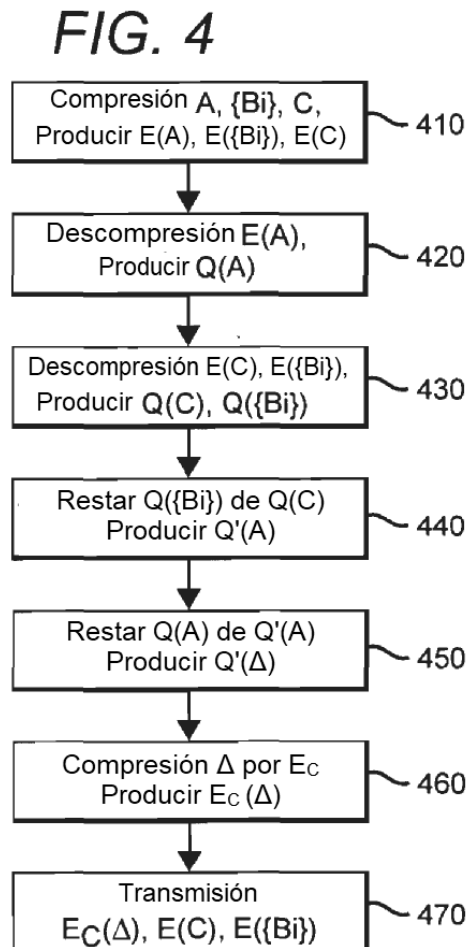


FIG. 4